

# SCORE Search Results Details for Application 10679580 and Search Result us-10-679-580a-1.rni.

[Score Home Page](#) [Retrieve Application List](#) [SCORE System Overview](#) [SCORE FAQ](#) [Comments / Suggestions](#)

This page gives you Search Results detail for the Application 10679580 and Search Result us-10-679-580a-1.rni.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 22, 2006, 22:47:34 ; Search time 2065 Seconds  
(without alignments)  
6110.779 Million cell updates/sec

Title: US-10-679-580A-1  
Perfect score: 6744  
Sequence: 1 ggacctcgcaggccaagaat.....aaaccaagaaatgcagcatt 6744

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/ina/1\_COMB.seq:\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/ina/5\_COMB.seq:\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6A\_COMB.seq:\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6B\_COMB.seq:\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7\_COMB.seq:\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/ina/H\_COMB.seq:\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PCTUS\_COMB.seq:\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PP\_COMB.seq:\*
- 9: /EMC\_Celerra\_SIDS3/ptodata/2/ina/RE\_COMB.seq:\*
- 10: /EMC\_Celerra\_SIDS3/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	1815.6	26.9	4173	4	US-09-512-581D-3	Sequence 3, Appli
2	1815.6	26.9	5271	4	US-09-512-581D-1	Sequence 1, Appli
3	1815.6	26.9	5355	4	US-09-512-581D-4	Sequence 4, Appli
4	1808	26.8	4418	3	US-09-949-016-4088	Sequence 4088, Ap
5	1686.6	25.0	1722	3	US-09-023-655-121	Sequence 121, App
6	437.6	6.5	452	3	US-09-513-999C-34334	Sequence 34334, A
7	206	3.1	1195	3	US-09-270-767-13950	Sequence 13950, A
8	161	2.4	295	3	US-09-513-999C-2815	Sequence 2815, Ap
9	149.4	2.2	351	3	US-09-471-276-620	Sequence 620, App
10	131	1.9	192506	3	US-09-949-016-15830	Sequence 15830, A
c 11	130.2	1.9	530	3	US-09-328-111-682	Sequence 682, App
12	88.4	1.3	601	3	US-09-949-016-144642	Sequence 144642,
13	86.4	1.3	601	3	US-09-949-016-144587	Sequence 144587,
c 14	80.2	1.2	612	3	US-09-902-540-1357	Sequence 1357, Ap
c 15	77.8	1.2	7218	2	US-08-232-463-14	Sequence 14, Appl
c 16	75.6	1.1	1039	3	US-09-902-540-1280	Sequence 1280, Ap
c 17	72.8	1.1	19124	2	US-08-487-826B-13	Sequence 13, Appl
18	72.6	1.1	1141	3	US-09-806-708B-22	Sequence 22, Appl
19	69.6	1.0	601	3	US-09-949-016-144594	Sequence 144594,
20	68	1.0	601	3	US-09-949-016-144593	Sequence 144593,
21	66	1.0	601	3	US-09-949-016-144592	Sequence 144592,
22	64.8	1.0	12313	3	US-09-949-016-13248	Sequence 13248, A
c 23	64.8	1.0	205044	3	US-09-949-016-15851	Sequence 15851, A
c 24	64.8	1.0	205044	3	US-09-949-016-15852	Sequence 15852, A
c 25	64.8	1.0	205044	3	US-09-949-016-15853	Sequence 15853, A
c 26	64.8	1.0	223471	3	US-09-949-016-12387	Sequence 12387, A
c 27	64.8	1.0	223471	3	US-09-949-016-12724	Sequence 12724, A
c 28	64.8	1.0	223471	3	US-09-949-016-12725	Sequence 12725, A
29	62.6	0.9	601	3	US-09-949-016-144622	Sequence 144622,
30	62.6	0.9	134987	3	US-09-949-016-15348	Sequence 15348, A
31	62.6	0.9	134987	3	US-09-949-016-15349	Sequence 15349, A
32	62.6	0.9	134987	3	US-09-949-016-15350	Sequence 15350, A
33	62.6	0.9	134987	3	US-09-949-016-15507	Sequence 15507, A
34	62.6	0.9	134987	3	US-09-949-016-15508	Sequence 15508, A
35	62.6	0.9	134987	3	US-09-949-016-15509	Sequence 15509, A
36	62	0.9	601	3	US-09-949-016-144624	Sequence 144624,
c 37	61.6	0.9	187169	3	US-09-949-016-12776	Sequence 12776, A
c 38	61.6	0.9	191569	3	US-09-949-016-15940	Sequence 15940, A
c 39	61.4	0.9	1141	3	US-09-806-708B-22	Sequence 22, Appl
c 40	60.4	0.9	114139	3	US-09-949-016-16536	Sequence 16536, A
41	57.8	0.9	931	3	US-09-270-767-13966	Sequence 13966, A
42	57.4	0.9	601	3	US-09-949-016-30531	Sequence 30531, A
43	57.4	0.9	601	3	US-09-949-016-37150	Sequence 37150, A
44	57.4	0.9	601	3	US-09-949-016-37164	Sequence 37164, A
45	57.4	0.9	601	3	US-09-949-016-145868	Sequence 145868,

## ALIGNMENTS

## RESULT 1

US-09-512-581D-3

; Sequence 3, Application US/09512581D

; Patent No. 6994992

; GENERAL INFORMATION:

; APPLICANT: Soto, Ana M.

; APPLICANT: Sonnenschein, Carlos

; APPLICANT: Geck, Peter

; APPLICANT: Szelei, Jozsef

; TITLE OF INVENTION: NOVEL ANDROGEN-INDUCED SUPPRESSOR OF CELL PROLIFERATION AND USES

```
; TITLE OF INVENTION:  THEREOF
; FILE REFERENCE:  34724-026
; CURRENT APPLICATION NUMBER:  US/09/512,581D
; CURRENT FILING DATE:   2000-02-24
; PRIOR APPLICATION NUMBER:  US 60/121,461
; PRIOR FILING DATE:  1999-02-24
; NUMBER OF SEQ ID NOS:  72
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 4173
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4173)
US-09-512-581D-3
```

```
Query Match          26.9%;  Score 1815.6;  DB 4;  Length 4173;
Best Local Similarity 69.7%;  Pred. No. 0;
Matches 2457;  Conservative 0;  Mismatches 1069;  Indels 0;  Gaps 0;
```

```
Qy      173  GGAAGATCGCTTACCCTCCGGGGGTAAAAGAGATCACCGACAAGATCACCACGGACGAGA 232
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      32  GAAAAATTACATATCCGCCTGGGGTCAAGGAAATATCAGATAAAATATCTAAAGAGGAGA 91

Qy      233  TGATCAAACGCCTGAAGATGGTAGTGAAAACCTTTATGGATATGGATCAGGACTCAGAAG 292
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      92  TGGTGAGACGATTAAAGATGGTTGTGAAAACCTTTATGGATATGGACCAGGACTCTGAAG 151

Qy      293  ATGAAAAACAGCAGTATCTCCCACTAGCCTTGCATCTTGCATCTGAATTCTTCCTCAGGA 352
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     152  AAGAAAAGGAGCTTTATTTAAACCTAGCTTTACATCTTGCTTCAGATTTTTTCTCAAGC 211

Qy      353  ACCCCAATAAAGATGTGCGTCTCCTTGTAGCATGTTGTTTGGCTGATATCTTTCGTATCT 412
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     212  ATCCTGGTAAAGATGTTGCTTACTGGTAGCCTGCTGCCTTGCTGATATTTTCAGGATTT 271

Qy      413  ATGCCCCAGAAGCTCCATATACTTCCCATGATAAACTTAAGGACATATTTTGTATTATTA 472
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     272  ATGCTCCTGAAGCTCCTTACACATCCCCTGATAAACTAAAGGATATATTTATGTTTATAA 331

Qy      473  CCAGACAATTAAAAGGTTTGGAGGATACAAAGAGTCCACAGTTTAATAGATACTTTTATT 532
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     332  CAAGACAGTTGAAGGGGCTAGAGGATACAAAGAGCCCAATTCAATAGGTATTTTTATT 391

Qy      533  TATTAGAGAATTTAGCTTGGGTAAATCATATAACATCTGCTTTGAATTGGAAGATTGCA 592
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     392  TACTTGAGAACATTGCTTGGGTCAAGTCATATAACATATGCTTTGAGTTAGAAGATAGCA 451

Qy      593  ATGAAATTTTATTTCAGCTTTTAGAACTCTCTTCTCAGTGATCAACAATAGCCACAATA 652
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     452  ATGAAATTTTCACCCAGCTATACAGAACCTTATTTTCAGTTATAACAATGGCCACAATC 511

Qy      653  AGAAGGTACAAATGCACATGCTAGATTTGATGAGTTCTATCATCATGGAAGGTGATGGAG 712
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     512  AGAAAGTCCATATGCACATGGTAGACCTTATGAGCTCTATATTTGTGAAGGTGATACAG 571

Qy      713  TTAAGGTAAGAAATTATTGGGCTCCATTCTTATTAACCTCATTCCTGCACATAAGAACTTAA 772
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     572  TGTCTCAGGAGCTTTTGGATACGGTTTTAGTAAATCTGGTACCTGCTCATAAGAATTTAA 631
```

# SCORE Search Results Details for Application 10679580 and Search Result us-10-679-580a-1.rge.

[Score Home Page](#) [Retrieve Application List](#) [SCORE System Overview](#) [SCORE FAQ](#) [Comments / Suggestions](#)

This page gives you Search Results detail for the Application 10679580 and Search Result us-10-679-580a-1.rge.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2006, 01:59:18 ; Search time 36612 Seconds  
(without alignments)  
11779.237 Million cell updates/sec

Title: US-10-679-580A-1  
Perfect score: 6744  
Sequence: 1 ggacctcgcaggccaagaat.....aaaccaagaaatgcagcatt 6744

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_env:\*  
2: gb\_pat:\*  
3: gb\_ph:\*  
4: gb\_pl:\*  
5: gb\_pr:\*  
6: gb\_ro:\*  
7: gb\_sts:\*  
8: gb\_sy:\*  
9: gb\_un:\*  
10: gb\_vi:\*  
11: gb\_ov:\*  
12: gb\_htg:\*  
13: gb\_in:\*  
14: gb\_om:\*  
15: gb\_ba:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description
	1	6744	100.0	6744	5	AF294791	AF294791 Homo sapi
	2	5177	76.8	5177	5	AB014548	AB014548 Homo sapi
	3	3744.4	55.5	3779	5	BC009650	BC009650 Homo sapi
	4	3145.2	46.6	4809	11	AJ851547	AJ851547 Gallus ga
c	5	2612.8	38.7	105001	5	AC105287	AC105287 Homo sapi
	6	2612.8	38.7	137481	12	AC027207	AC027207 Homo sapi
c	7	2612.8	38.7	162775	12	AC023006	AC023006 Homo sapi
	8	2527	37.5	3305	2	CQ727538	CQ727538 Sequence
	9	2355.6	34.9	4654	11	BC098992	BC098992 Xenopus l
	10	2353.8	34.9	4328	11	AY695731	AY695731 Xenopus l
	11	2291	34.0	160962	12	AC022862	AC022862 Homo sapi
	12	2198	32.6	2212	2	BD159124	BD159124 Primer fo
	13	2198	32.6	2212	2	AX881562	AX881562 Sequence
	14	2198	32.6	2212	5	AK021757	AK021757 Homo sapi
	15	2134.2	31.6	5275	6	AK172992	AK172992 Mus muscu
	16	1983.2	29.4	2002	2	BD156759	BD156759 Primer fo
	17	1983.2	29.4	2002	2	AX877628	AX877628 Sequence
	18	1983.2	29.4	2002	5	AK001828	AK001828 Homo sapi
	19	1822	27.0	5309	5	AB023196	AB023196 Homo sapi
	20	1822	27.0	7444	5	HS26H231	AL137201 Novel hum
	21	1822	27.0	7473	2	CQ490556	CQ490556 Sequence
	22	1822	27.0	7473	2	CQ493591	CQ493591 Sequence
	23	1822	27.0	7473	2	CQ496390	CQ496390 Sequence
	24	1820.4	27.0	5297	11	BC086289	BC086289 Xenopus l
	25	1817.2	26.9	5253	5	HSU95825	U95825 Human andro
	26	1808.8	26.8	5276	11	AJ851551	AJ851551 Gallus ga
	27	1789	26.5	5317	6	AY102267	AY102267 Mus muscu
	28	1785.4	26.5	5472	6	BC096539	BC096539 Mus muscu
	29	1773.2	26.3	5154	11	BC100220	BC100220 Xenopus l
	30	1706.2	25.3	2265	2	CS169195	CS169195 Sequence
	31	1706.2	25.3	2265	2	AX835316	AX835316 Sequence
	32	1706.2	25.3	2265	5	AK098331	AK098331 Homo sapi
	33	1706.2	25.3	2705	5	BC041361	BC041361 Homo sapi
	34	1703	25.3	1936	5	AY550968	AY550968 Homo sapi
	35	1686.6	25.0	1722	2	AR379576	AR379576 Sequence
	36	1594.4	23.6	1976	6	BC021408	BC021408 Mus muscu
c	37	1489.6	22.1	188061	12	AC173425	AC173425 Bos tauru
	38	1489	22.1	1489	2	BD160058	BD160058 Primer fo
	39	1489	22.1	1489	2	AX882999	AX882999 Sequence
	40	1489	22.1	1489	5	AK023592	AK023592 Homo sapi
	41	1357.8	20.1	5944	6	AK122414	AK122414 Mus muscu
	42	979.2	14.5	3632	11	AY695732	AY695732 Xenopus l
	43	918	13.6	1852	5	HSU50533	U50533 Human BRCA2
	44	806	12.0	4252	5	BC039256	BC039256 Homo sapi
c	45	761.2	11.3	110000	12	AC156034_2	Continuation (3 of

## ALIGNMENTS

## RESULT 1

AF294791

LOCUS AF294791 6744 bp mRNA linear PRI 22-MAR-2004

DEFINITION Homo sapiens SCC-112 (SCC-112) mRNA, complete cds.

ACCESSION AF294791

VERSION AF294791.1 GI:21951801  
 KEYWORDS .  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 6744)  
 AUTHORS Kumar,D., Sakabe,I., Patel,S., Zhang,Y., Ahmad,I., Gehan,E.A.,  
 Whiteside,T.L. and Kasid,U.  
 TITLE SCC-112, a novel cell cycle-regulated molecule, exhibits reduced  
 expression in human renal carcinomas  
 JOURNAL Gene 328, 187-196 (2004)  
 PUBMED 15019998  
 REFERENCE 2 (bases 1 to 6744)  
 AUTHORS Kumar,D. and Kasid,U.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-AUG-2000) Radiation, Medicine, 3970 Reservoir Rd, NW,  
 Washington, DC 20007, USA

FEATURES Location/Qualifiers  
 source 1. .6744  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="4"  
 gene 1. .6744  
 /gene="SCC-112"  
 CDS 232. .4125  
 /gene="SCC-112"  
 /codon\_start=1  
 /product="SCC-112"  
 /protein\_id="AAM82347.1"  
 /db\_xref="GI:21951802"  
 /translation="MIKRLKMVVKTFMDMDQDSEDEKQQYLPLALHLASEFFLRNP  
 DVRLLVACCLADIFRIYAPEAPYTSKDKLKDIFLFIITRQLKGLEDTKSPQFNRYFYLL  
 ENLAWKSYNICFELEDCEIFIQLFRTLFVSVINNSHNKKVQMHMLDLMSSIIMEGDG  
 VTQELLGSLINLI PAHKNLNKQSFDAKVLKRTVQTI EACIANFFNQVLVLRSSV  
 SDLSEHVFDLIQELFAIDPHLLSVMPQLEFKLSNDGEERLAVVRLAKLFGSKDS  
 LATQNRPLWQCFLGRFNDIHVPVRLESVKFASHCLMNHPDLAKDLTEYLKVRSHDPEE  
 AIRHDVIVTIITAARKDLALVNDQLLGFVRERTLDKRWRVRKEAMMGLAQLYKKYCLH  
 GEAGKEAAEKVSWIKDKLLHIYYQNSIDDKLLVEKIFAQYLVPHNLETEERMKCLYYL  
 YASLDPNAVKALNEMWKQCNMLRSHVRELLDLHKQPTSEANCSAMFGKLMITAKNLPD  
 PGKAQDFVKKFNQVLGDDKLRSLLELLISPTCSCKQADICVREIARKLANPKQPTNP  
 FLEMVKFLLERIAPVHIDSEASALVKLMNKSI EGTADDEEEGVSPDTAIRSGLELLK  
 VLSFTHPTSFHSAETYESLLQCLRMEDDKVAEAAIQIFRNTGHKIEIDLPIRSTLIP  
 ILHQAKRGT PHQAKQAVHCIHAIFTNKEVQLAQIFEPLSRSLNADVPEQLITPLVSL  
 GHISMLAPDQFASPMKSVVANFIVKDLLMNDRSTGEKNGKLWSPDEEVSPEVLAKVQA  
 IKLLVRWLLGMKNQSKSANSTLRLLSAMLVSEGLTEQKRISKSDMSRLRLAAGSAI  
 MKLAQEPYHEIITPEQFQLCALVINDECYQVRQIFAQKLHKALVKLLLPLEYMAIFA  
 LCAKDPVKERRAHARQCLLNISIRREYIKQNPMA TEKLLSLLPEYVVPYMIHLAHD  
 PDFTRSQDQDLRDIKECLWFMLEVLMTKNENNSHAFMKMAENIKLTRDAQSPDESK  
 TNEKLYTVCDVALCVINSKALCNADSPKDPVLPKFFTPQEKDFCNDKSYISEETRV  
 LLLTGKPKPAGVLGAVNKPLSATGRKPYVRSTGTETGSGNINVNSELPSTGNRSREQS  
 SEAAETGVSENEENPVRIISVTPVKNIDPVKNKEINSQATQGNISDRGKKRTVTAA  
 GAENIQKKTDEKVDSESGPPAPSKPRRGRRPKSESQGNATKNDDL NKPINKGRKRAAVG  
 QESPGGLEAGNAKAPKLQDLAKKAAPAERQIDLQR"

## ORIGIN

Query Match 100.0%; Score 6744; DB 5; Length 6744;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 6744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGACCTCGCAGGCCAAGAATTCGGCACGAGGGGCGCCGGCTCCCGGGGCACGGACGGCCG	60
Db	1	GGACCTCGCAGGCCAAGAATTCGGCACGAGGGGCGCCGGCTCCCGGGGCACGGACGGCCG	60
Qy	61	GGCGCGCGCCTCTGCGAGGGGCGTCCGGGTCCGAGTCGGCGGTCCGGGCCGGCGCGAGGT	120
Db	61	GGCGCGCGCCTCTGCGAGGGGCGTCCGGGTCCGAGTCGGCGGTCCGGGCCGGCGCGAGGT	120
Qy	121	GCGTGCGGGCGGGCCGCGGGGGTCCGGACGGACACAAGCGCACACACTCCCGGAAGATC	180
Db	121	GCGTGCGGGCGGGCCGCGGGGGTCCGGACGGACACAAGCGCACACACTCCCGGAAGATC	180
Qy	181	GCTTACCCTCCGGGGGTAAAAGAGATCACCACAAGATCACCACGGACGAGATGATCAA	240
Db	181	GCTTACCCTCCGGGGGTAAAAGAGATCACCACAAGATCACCACGGACGAGATGATCAA	240
Qy	241	CGCCTGAAGATGGTAGTGAAAACCTTTATGGATATGGATCAGGACTCAGAAGATGAAAA	300
Db	241	CGCCTGAAGATGGTAGTGAAAACCTTTATGGATATGGATCAGGACTCAGAAGATGAAAA	300
Qy	301	CAGCAGTATCTCCCACTAGCCTTGATCTTGATCTGAATTCTTCCTCAGGAACCCCAAT	360
Db	301	CAGCAGTATCTCCCACTAGCCTTGATCTTGATCTGAATTCTTCCTCAGGAACCCCAAT	360
Qy	361	AAAGATGTGCGTCTCCTTGATGATGTTGTTGGCTGATATCTTTCGTATCTATGCCCCA	420
Db	361	AAAGATGTGCGTCTCCTTGATGATGTTGTTGGCTGATATCTTTCGTATCTATGCCCCA	420
Qy	421	GAAGCTCCATATACTTCCCATGATAAACTTAAGGACATATTTTGTATTACCAGACAA	480
Db	421	GAAGCTCCATATACTTCCCATGATAAACTTAAGGACATATTTTGTATTACCAGACAA	480
Qy	481	TTAAAAGGTTTGGAGGATACAAAGAGTCCACAGTTAATAGATACTTTATTTATTAGAG	540
Db	481	TTAAAAGGTTTGGAGGATACAAAGAGTCCACAGTTAATAGATACTTTATTTATTAGAG	540
Qy	541	AATTTAGCTTGGGTAAATCATATAACATCTGCTTTGAATTGGAAGATTGCAATGAAATT	600
Db	541	AATTTAGCTTGGGTAAATCATATAACATCTGCTTTGAATTGGAAGATTGCAATGAAATT	600
Qy	601	TTTATTCAGCTTTTGTAGAACTCTCTCTCAGTGATCAACAATAGCCACAATAAGAAGGTA	660
Db	601	TTTATTCAGCTTTTGTAGAACTCTCTCTCAGTGATCAACAATAGCCACAATAAGAAGGTA	660
Qy	661	CAAATGCACATGCTAGATTTGATGAGTTCTATCATCATGGAAGGTGATGGAGTTACTCAA	720
Db	661	CAAATGCACATGCTAGATTTGATGAGTTCTATCATCATGGAAGGTGATGGAGTTACTCAA	720
Qy	721	GAATTATTGGGCTCCATTCTTATTAACCTCATTCCTGCACATAAGAACTTAAATAAACAG	780
Db	721	GAATTATTGGGCTCCATTCTTATTAACCTCATTCCTGCACATAAGAACTTAAATAAACAG	780
Qy	781	TCCTTTGACCTTGCAAAGTGCTATTGAAAAGAACAGTCCAGACTATTGAGGCATGCATT	840
Db	781	TCCTTTGACCTTGCAAAGTGCTATTGAAAAGAACAGTCCAGACTATTGAGGCATGCATT	840
Qy	841	GCCAATTTTTCATCAAGTCCTGGTGCTGGGAAGATCATCAGTAAGTGATTGTCAGAA	900
Db	841	GCCAATTTTTCATCAAGTCCTGGTGCTGGGAAGATCATCAGTAAGTGATTGTCAGAA	900

Qy	901	CATGTATTTGATCTGATTCAAGAACTTTTGGCTATAGATCCTCATTTATTATTATCCGTC	960
Db	901	CATGTATTTGATCTGATTCAAGAACTTTTGGCTATAGATCCTCATTTATTATTATCCGTC	960
Qy	961	ATGCCACAGCTTGAATTCAACTAAAGAGCAATGATGGAGAAGAGCGATTAGCTGTTGTT	1020
Db	961	ATGCCACAGCTTGAATTCAACTAAAGAGCAATGATGGAGAAGAGCGATTAGCTGTTGTT	1020
Qy	1021	CGACTTCTAGCTAAATTGTTTGGCTCCAAAGATTCTGATTTGGCAACACAGAATCGTCCT	1080
Db	1021	CGACTTCTAGCTAAATTGTTTGGCTCCAAAGATTCTGATTTGGCAACACAGAATCGTCCT	1080
Qy	1081	CTTTGGCAATGTTTCTTGACGATTTAATGATATTCATGTTCTGTGAGATTAGAAAGT	1140
Db	1081	CTTTGGCAATGTTTCTTGACGATTTAATGATATTCATGTTCTGTGAGATTAGAAAGT	1140
Qy	1141	GTGAAATTTGCCAGTCATTGTTAATGAATCACCAGATTTAGCGAAGGATCTCACAGAA	1200
Db	1141	GTGAAATTTGCCAGTCATTGTTAATGAATCACCAGATTTAGCGAAGGATCTCACAGAA	1200
Qy	1201	TATTTAAAGGTTAGATCACATGATCCAGAAGAAGCTATTCGTATGATGTCATTGTTACT	1260
Db	1201	TATTTAAAGGTTAGATCACATGATCCAGAAGAAGCTATTCGTATGATGTCATTGTTACT	1260
Qy	1261	ATAATAACAGCTGCCAAGAGGGACCTGGCCTTAGTAAATGATCAGCTGCTTGGCTTTGTA	1320
Db	1261	ATAATAACAGCTGCCAAGAGGGACCTGGCCTTAGTAAATGATCAGCTGCTTGGCTTTGTA	1320
Qy	1321	AGGGAAAGAACACTGGATAAAACGGTGGCGAGTAAGAAAAGAAGCTATGATGGGTCTGGCT	1380
Db	1321	AGGGAAAGAACACTGGATAAAACGGTGGCGAGTAAGAAAAGAAGCTATGATGGGTCTGGCT	1380
Qy	1381	CAGCTTTATAAGAAATACTGTCTTCATGGTGAAGCAGGAAAGGAAGCTGCAGAGAAAGTC	1440
Db	1381	CAGCTTTATAAGAAATACTGTCTTCATGGTGAAGCAGGAAAGGAAGCTGCAGAGAAAGTC	1440
Qy	1441	AGCTGGATAAAAGGACAAACTTCTGCATATTTATTATCAGAACAGCATTGACGACAAACTG	1500
Db	1441	AGCTGGATAAAAGGACAAACTTCTGCATATTTATTATCAGAACAGCATTGACGACAAACTG	1500
Qy	1501	TTGGTAGAGAAAATCTTTGCTCAGTATCTTGTCCCCACAACCTGGAAACAGAAGAGAGA	1560
Db	1501	TTGGTAGAGAAAATCTTTGCTCAGTATCTTGTCCCCACAACCTGGAAACAGAAGAGAGA	1560
Qy	1561	ATGAAATGCTTATATTACTTATATGCTAGTTTGGATCCAAATGCTGTAAAAGCTCTCAAC	1620
Db	1561	ATGAAATGCTTATATTACTTATATGCTAGTTTGGATCCAAATGCTGTAAAAGCTCTCAAC	1620
Qy	1621	GAAATGTGGAAGTGTGAGAACATGCTTCGGAGCCATGTACGCGAACTATTGGATTTCAC	1680
Db	1621	GAAATGTGGAAGTGTGAGAACATGCTTCGGAGCCATGTACGCGAACTATTGGATTTCAC	1680
Qy	1681	AAGCAGCCTACATCAGAGGCTAACTGTTCTGCCATGTTTGGAAAAGCTGATGACCATAGCA	1740
Db	1681	AAGCAGCCTACATCAGAGGCTAACTGTTCTGCCATGTTTGGAAAAGCTGATGACCATAGCA	1740
Qy	1741	AAGAATTTGCCTGACCCCGGGAAGCACAAGATTTTGTGAAGAAATTAACCAGGTTCTC	1800
Db	1741	AAGAATTTGCCTGACCCCGGGAAGCACAAGATTTTGTGAAGAAATTAACCAGGTTCTC	1800
Qy	1801	GGCGATGATGAGAACTTCGGTCTCAGTTGGAGTTATTAATTAGCCCAACCTGTTCTTGC	1860



# SCORE Search Results Details for Application 10679580 and Search Result us-10-679-580a-1.rng.

[Score Home Page](#) [Retrieve Application List](#) [SCORE System Overview](#) [SCORE FAQ](#) [Comments / Suggestions](#)

This page gives you Search Results detail for the Application 10679580 and Search Result us-10-679-580a-1.rng.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2006, 00:55:58 ; Search time 3752 Seconds  
(without alignments)  
12532.208 Million cell updates/sec

Title: US-10-679-580A-1  
Perfect score: 6744  
Sequence: 1 ggacctcgcaggccaagaat.....aaaccaagaaatgcagcatt 6744

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_8:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*  
14: geneseqn2005s:\*  
15: geneseqn2006s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Score	% Query Match	Length	DB	ID	Description
No.						
1	6744	100.0	6744	8	ABX14056	Abx14056 cDNA enco
2	6744	100.0	6744	14	ADX06262	Adx06262 Cyclin-de
3	6555.6	97.2	6956	8	ACC72031	Acc72031 BCU0720 g
4	5385.8	79.9	5469	6	ABQ99271	Abq99271 Human cod
5	5177	76.8	5177	14	AED18329	Aed18329 Fibrotic
6	5177	76.8	5177	15	AEF74956	Aef74956 Human pol
7	3875.4	57.5	3957	3	AAA47423	Aaa47423 Sequence
8	2198	32.6	2212	4	AAH17132	Aah17132 Human cDN
9	2128.4	31.6	2295	4	AAS02396	Aas02396 Human sec
10	1983.2	29.4	2002	4	AAH14767	Aah14767 Human cDN
11	1953.4	29.0	2496	6	ABK35358	Abk35358 Human cDN
12	1823.6	27.0	7473	5	ABV25469	Abv25469 Human pro
13	1822	27.0	5309	15	AEF74563	Aef74563 Human pol
14	1822	27.0	7444	13	ADR25694	Adr25694 Breast ca
15	1822	27.0	7473	5	ABV22430	Abv22430 Human pro
16	1822	27.0	7473	5	ABV28244	Abv28244 Human pro
17	1815.6	26.9	5271	3	AAA28051	Aaa28051 Human and
18	1815.6	26.9	5355	3	AAA28052	Aaa28052 Human and
19	1706.2	25.3	2265	11	ADM03755	Adm03755 Human cDN
20	1706.2	25.3	2265	14	AEC86685	Aec86685 Human cDN
21	1686.6	25.0	1722	11	ADI30795	Adi30795 Human cDN
22	1686.6	25.0	1722	13	ADS82862	Ads82862 Human lym
23	1489	22.1	1489	4	AAH18066	Aah18066 Human cDN
24	1148.8	17.0	1217	3	AAF18338	Aaf18338 Lung canc
25	726.8	10.8	785	4	AAH07766	Aah07766 Human cDN
26	686.4	10.2	1104	6	ABS56684	Abs56684 Euchromos
27	670.4	9.9	4128	13	ADR08232	Adr08232 Full leng
28	666.8	9.9	772	4	AAH03894	Aah03894 Human cDN
29	593.2	8.8	663	4	AAH07307	Aah07307 Human cDN
c 30	581.4	8.6	584	6	ABK44862	Abk44862 cDNA enco
c 31	478.6	7.1	542	4	AAH09415	Aah09415 Human cDN
c 32	462.2	6.9	512	4	AAH12296	Aah12296 Human cDN
33	437.6	6.5	452	3	AAC30259	Aac30259 Human sec
c 34	430.8	6.4	495	4	AAH12739	Aah12739 Human cDN
c 35	429.8	6.4	471	9	ACH16302	Ach16302 Human adu
36	417.2	6.2	434	8	ABZ17884	Abz17884 S2 subtra
37	408	6.0	408	6	ABV95661	Abv95661 Human pan
c 38	389.2	5.8	391	6	ABL38139	Abl38139 Human col
39	381	5.6	413	9	ACH49024	Ach49024 Human leu
40	376.8	5.6	422	9	ACH48906	Ach48906 Human leu
c 41	360.2	5.3	376	6	ABL37597	Abl37597 Human col
42	359	5.3	2416	12	ADO35588	Ado35588 Novel mou
c 43	346.2	5.1	706	10	ADD35009	Add35009 Mouse mit
44	335.2	5.0	380	14	AEE11851	Aee11851 Hamster c
45	335.2	5.0	380	14	AEE15493	Aee15493 Hamster S

## ALIGNMENTS

## RESULT 1

ABX14056

ID ABX14056 standard; cDNA; 6744 BP.

XX

AC ABX14056;

XX  
DT 24-FEB-2003 (first entry)  
XX  
DE cDNA encoding human SCC-112 tumour supressor gene.  
XX  
KW ss; gene; human; chromosome 4p14; SCC-112; cancer; apoptosis; allodynia;  
KW degenerative disorder; metastasis inhibition; breast cancer; causalgia;  
KW kidney cancer; bladder cancer; pancreatic cancer; colon cancer;  
KW squamous cell carcinoma; head trauma; spinal cord injury; herpes zoster;  
KW global and focal ischaemic and haemorrhagic stroke; epilepsy; neuralgia;  
KW hypoxia-induced nerve cell damage; anxiety; diabetes mellitus;  
KW cardiac arrest; spinal cord lesion; stomach cancer; lung cancer;  
KW neonatal distress; Alzheimer's disease; uterine cancer; colon cancer;  
KW multiple sclerosis; phantom limb pain; hyperalgesia; Down's syndrome;  
KW Huntington's disease; Parkinson's disease; Korsakoff's syndrome;  
KW amyotrophic lateral sclerosis; cell survival; cell proliferation;  
KW tumour supressor.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 232. .4125  
FT /\*tag= a  
FT /product= "Human SCC-112"  
FT polyA\_signal 6721. .6726  
FT /\*tag= b  
FT /standard\_name= "PolyA signal"  
XX  
PN WO200281641-A2.  
XX  
PD 17-OCT-2002.  
XX  
PF 08-APR-2002; 2002WO-US010850.  
XX  
PR 06-APR-2001; 2001US-0281780P.  
XX  
PA (GEOU ) UNIV GEORGETOWN.  
XX  
PI Kasid UN, Kumar D, Ahmad I;  
XX  
DR WPI; 2003-103330/09.  
DR P-PSDB; ABG72803.  
XX  
PT New isolated SSC (undefined) tumor suppressor polypeptides and  
PT polynucleotides, useful for diagnosing, preventing or treating cancer or  
PT degenerative disease, e.g. Alzheimer's Disease, Huntington's disease, or  
PT multiple sclerosis.  
XX  
PS Claim 1; Fig 1B; 83pp; English.  
XX  
CC The invention relates to a new isolated polypeptide SCC-112. The SCC-112  
CC polypeptides and polynucleotides are useful for diagnosing, preventing or  
CC treating cancer (e.g. breast cancer, kidney cancer, bladder cancer,  
CC pancreatic cancer, colon cancer, squamous cell carcinoma, uterine cancer,  
CC stomach cancer, colon cancer, lung cancer); or degenerative disease or  
CC disorder (e.g. global and focal ischaemic and haemorrhagic stroke, head  
CC trauma, spinal cord injury, hypoxia-induced nerve cell damage, nerve cell  
CC damage caused by cardiac arrest or neonatal distress, epilepsy, anxiety,  
CC diabetes mellitus, multiple sclerosis, phantom limb pain, causalgia,  
CC neuralgias, herpes zoster, spinal cord lesions, hyperalgesia, allodynia,  
CC Alzheimer's disease, Huntington's disease, Parkinson's disease,

CC amyotrophic lateral sclerosis, Down's syndrome and Korsakoff's syndrome).  
CC The polypeptides and polynucleotides are also useful for inducing  
CC apoptosis in cancer cells, increasing survival or proliferation of a  
CC cell, or inhibiting cancer cell proliferation and/or metastasis in a  
CC cancer patient. The polynucleotides can be used as probes to detect  
CC complementary nucleotide sequences, or as primers to obtain additional  
CC copies of the polynucleotides. SSC-112 may also be used for identifying  
CC drugs for treatment of cancers. The present sequence represents cDNA  
CC encoding the human SCC-112 tumour supressor gene which is located on  
CC chromosome 4p14

XX  
SQ Sequence 6744 BP; 2156 A; 1243 C; 1400 G; 1945 T; 0 U; 0 Other;

Query Match 100.0%; Score 6744; DB 8; Length 6744;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 6744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGACCTCGCAGGCCAAGAATTCGGCACGAGGGGCGCCGGCTCCCGGGGCACGGACGGCCG	60
Db	1	GGACCTCGCAGGCCAAGAATTCGGCACGAGGGGCGCCGGCTCCCGGGGCACGGACGGCCG	60
Qy	61	GGCGCGCGCCTCTGCGAGGGGCGTCCGGTCCGAGTCGGCGGTCCGGGCCGGCGCGAGGT	120
Db	61	GGCGCGCGCCTCTGCGAGGGGCGTCCGGTCCGAGTCGGCGGTCCGGGCCGGCGCGAGGT	120
Qy	121	GCGTGCGGGCGGGCCGCGGGGGTCCCGACGGACACAAGCGCACACACTCCCGGAAGATC	180
Db	121	GCGTGCGGGCGGGCCGCGGGGGTCCCGACGGACACAAGCGCACACACTCCCGGAAGATC	180
Qy	181	GCTTACCCTCCGGGGGTAAAAGAGATCACCGACAAGATCACACGGACGAGATGATCAAA	240
Db	181	GCTTACCCTCCGGGGGTAAAAGAGATCACCGACAAGATCACACGGACGAGATGATCAAA	240
Qy	241	CGCCTGAAGATGGTAGTGAAAACCTTTATGGATATGGATCAGGACTCAGAAGATGAAAAA	300
Db	241	CGCCTGAAGATGGTAGTGAAAACCTTTATGGATATGGATCAGGACTCAGAAGATGAAAAA	300
Qy	301	CAGCAGTATCTCCCACTAGCCTTGATCTTGCATCTGAATTCTTCCTCAGGAACCCCAAT	360
Db	301	CAGCAGTATCTCCCACTAGCCTTGATCTTGCATCTGAATTCTTCCTCAGGAACCCCAAT	360
Qy	361	AAAGATGTGCGTCTCCTTGATGATGTTGTTGGCTGATATCTTTCGTATCTATGCCCCA	420
Db	361	AAAGATGTGCGTCTCCTTGATGATGTTGTTGGCTGATATCTTTCGTATCTATGCCCCA	420
Qy	421	GAAGCTCCATATACTTCCCATGATAAACTTAAGGACATATTTTGTATTATTACCAGACAA	480
Db	421	GAAGCTCCATATACTTCCCATGATAAACTTAAGGACATATTTTGTATTATTACCAGACAA	480
Qy	481	TTAAAAGGTTTGGAGGATACAAAGAGTCCACAGTTTAATAGATACTTTATTTATTAGAG	540
Db	481	TTAAAAGGTTTGGAGGATACAAAGAGTCCACAGTTTAATAGATACTTTATTTATTAGAG	540
Qy	541	AATTTAGCTTGGGTAAATCATATAACATCTGCTTTGAATTGGAAGATTGCAATGAAATT	600
Db	541	AATTTAGCTTGGGTAAATCATATAACATCTGCTTTGAATTGGAAGATTGCAATGAAATT	600
Qy	601	TTTATTCAGCTTTTGTAGAACTCTCTCTCAGTGATCAACAATAGCCACAATAAGAAGGTA	660
Db	601	TTTATTCAGCTTTTGTAGAACTCTCTCTCAGTGATCAACAATAGCCACAATAAGAAGGTA	660

Qy	661	CAAATGCACATGCTAGATTGATGAGTTCTATCATCATGGAAGGTGATGGAGTTACTCAA	720
Db	661	CAAATGCACATGCTAGATTGATGAGTTCTATCATCATGGAAGGTGATGGAGTTACTCAA	720
Qy	721	GAATTATTGGGCTCCATTCTTATTAACCTCATTCCTGCACATAAGAACTTAAATAAACAG	780
Db	721	GAATTATTGGGCTCCATTCTTATTAACCTCATTCCTGCACATAAGAACTTAAATAAACAG	780
Qy	781	TCCTTTGACCTTGCAAAAGTGCTATTGAAAAGAACAGTCCAGACTATTGAGGCATGCATT	840
Db	781	TCCTTTGACCTTGCAAAAGTGCTATTGAAAAGAACAGTCCAGACTATTGAGGCATGCATT	840
Qy	841	GCCAAATTTTTCAATCAAGTCCTGGTGCTGGGAAGATCATCAGTAAGTGATTGTCAGAA	900
Db	841	GCCAAATTTTTCAATCAAGTCCTGGTGCTGGGAAGATCATCAGTAAGTGATTGTCAGAA	900
Qy	901	CATGTATTTGATCTGATTGAGAACTTTTTGCTATAGATCCTCATTTATTATTATCCGTC	960
Db	901	CATGTATTTGATCTGATTGAGAACTTTTTGCTATAGATCCTCATTTATTATTATCCGTC	960
Qy	961	ATGCCACAGCTTGAATTCAAATAAGAGCAATGATGGAGAAGAGCGATTAGCTGTTGTT	1020
Db	961	ATGCCACAGCTTGAATTCAAATAAGAGCAATGATGGAGAAGAGCGATTAGCTGTTGTT	1020
Qy	1021	CGACTTCTAGCTAAATTGTTTGGCTCCAAAGATTCTGATTGGCAACACAGAATCGTCCT	1080
Db	1021	CGACTTCTAGCTAAATTGTTTGGCTCCAAAGATTCTGATTGGCAACACAGAATCGTCCT	1080
Qy	1081	CTTTGGCAATGTTTCTTGGACGATTTAATGATATTCATGTTCTGTGAGATTAGAAAGT	1140
Db	1081	CTTTGGCAATGTTTCTTGGACGATTTAATGATATTCATGTTCTGTGAGATTAGAAAGT	1140
Qy	1141	GTGAAATTTGCCAGTCATTGTTTAATGAATCACCCAGATTTAGCGAAGGATCTCACAGAA	1200
Db	1141	GTGAAATTTGCCAGTCATTGTTTAATGAATCACCCAGATTTAGCGAAGGATCTCACAGAA	1200
Qy	1201	TATTTAAAGGTTAGATCACATGATCCAGAAGAAGCTATTCGTATGATGTCATTGTTACT	1260
Db	1201	TATTTAAAGGTTAGATCACATGATCCAGAAGAAGCTATTCGTATGATGTCATTGTTACT	1260
Qy	1261	ATAATAACAGCTGCCAAGAGGGACCTGGCCTTAGTAAATGATCAGCTGCTTGGCTTTGTA	1320
Db	1261	ATAATAACAGCTGCCAAGAGGGACCTGGCCTTAGTAAATGATCAGCTGCTTGGCTTTGTA	1320
Qy	1321	AGGGAAAGAACACTGGATAAACCGGTGGCGAGTAAGAAAAGAAGCTATGATGGGTCTGGCT	1380
Db	1321	AGGGAAAGAACACTGGATAAACCGGTGGCGAGTAAGAAAAGAAGCTATGATGGGTCTGGCT	1380
Qy	1381	CAGCTTTATAAGAAATACTGTCTTCATGGTGAAGCAGGAAAGGAAGCTGCAGAGAAAGTC	1440
Db	1381	CAGCTTTATAAGAAATACTGTCTTCATGGTGAAGCAGGAAAGGAAGCTGCAGAGAAAGTC	1440
Qy	1441	AGCTGGATAAAGGACAAACTTCTGCATATTTATTATCAGAACAGCATTGACGACAAACTG	1500
Db	1441	AGCTGGATAAAGGACAAACTTCTGCATATTTATTATCAGAACAGCATTGACGACAAACTG	1500
Qy	1501	TTGGTAGAGAAAATCTTTGCTCAGTATCTTGTCCCCCAACCTGGAAACAGAAGAGAGA	1560
Db	1501	TTGGTAGAGAAAATCTTTGCTCAGTATCTTGTCCCCCAACCTGGAAACAGAAGAGAGA	1560
Qy	1561	ATGAAATGCTTATATTACTTATATGCTAGTTTGGATCCAAATGCTGTAAAAGCTCTCAAC	1620

Db	1561	 ATGAAATGCTTATATTACTTATATGCTAGTTTGGATCCAAATGCTGTAAAAGCTCTCAAC	1620
Qy	1621	GAAATGTGGAAGTGTGAGAACATGCTTCGGAGCCATGTACGCGAACTATTGGATTTCAC	1680
Db	1621	 GAAATGTGGAAGTGTGAGAACATGCTTCGGAGCCATGTACGCGAACTATTGGATTTCAC	1680
Qy	1681	AAGCAGCCTACATCAGAGGCTAACTGTTCTGCCATGTTTGAAAAGTATGACCATAGCA	1740
Db	1681	 AAGCAGCCTACATCAGAGGCTAACTGTTCTGCCATGTTTGAAAAGTATGACCATAGCA	1740
Qy	1741	AAGAATTTGCCTGACCCCGGAAAGCACAAGATTTTGTGAAGAAATTTAACCAGGTTCTC	1800
Db	1741	 AAGAATTTGCCTGACCCCGGAAAGCACAAGATTTTGTGAAGAAATTTAACCAGGTTCTC	1800
Qy	1801	GGCGATGATGAGAACTTCGGTCTCAGTTGGAGTTATTAATTAGCCCAACCTGTTCTTGC	1860
Db	1801	 GGCGATGATGAGAACTTCGGTCTCAGTTGGAGTTATTAATTAGCCCAACCTGTTCTTGC	1860
Qy	1861	AAACAAGCAGATATTTGTGTGAGAGAAATAGCCCGAACTTGCAAATCCTAAGCAACCA	1920
Db	1861	 AAACAAGCAGATATTTGTGTGAGAGAAATAGCCCGAACTTGCAAATCCTAAGCAACCA	1920
Qy	1921	ACAAATCCTTTTCTAGAGATGGTCAAATTTCTGTTGAAAGAATCGACCTGTGCACATT	1980
Db	1921	 ACAAATCCTTTTCTAGAGATGGTCAAATTTCTGTTGAAAGAATCGACCTGTGCACATT	1980
Qy	1981	GATTCAGAAGCCATAAGTGCACTAGTGAATTTGATGAATAAGTCAATAGAGGGGACAGCA	2040
Db	1981	 GATTCAGAAGCCATAAGTGCACTAGTGAATTTGATGAATAAGTCAATAGAGGGGACAGCA	2040
Qy	2041	GATGATGAAGAGGAGGGTGTAAGTCCAGATACAGCTATCCGTTGAGGACTTGAACCTCTT	2100
Db	2041	 GATGATGAAGAGGAGGGTGTAAGTCCAGATACAGCTATCCGTTGAGGACTTGAACCTCTT	2100
Qy	2101	AAGGTTCTGTCTTTTACACATCCTACCTCGTTCCACTCTGCAGAGACATATGAGTCCTTG	2160
Db	2101	 AAGGTTCTGTCTTTTACACATCCTACCTCGTTCCACTCTGCAGAGACATATGAGTCCTTG	2160
Qy	2161	TTACAGTGCCTAAGAATGGAGGATGACAAGGTAGCAGAAGCTGCTATTCAAATTTTGA	2220
Db	2161	 TTACAGTGCCTAAGAATGGAGGATGACAAGGTAGCAGAAGCTGCTATTCAAATTTTGA	2220
Qy	2221	AATACAGGTCAAAAATAGAAACAGACCTTCCCAGATACGATCGACCTTAATCCCATT	2280
Db	2221	 AATACAGGTCAAAAATAGAAACAGACCTTCCCAGATACGATCGACCTTAATCCCATT	2280
Qy	2281	TTACATCAAAAAGCAAAGAGGGGTACTCCACACCAAGCAAAACAGGCTGTGCACTGTATA	2340
Db	2281	 TTACATCAAAAAGCAAAGAGGGGTACTCCACACCAAGCAAAACAGGCTGTGCACTGTATA	2340
Qy	2341	CACGCCATATTCACAAATAAAGAAGTCCAGCTTGACAGATTTTGTAGCCACTCAGTAGG	2400
Db	2341	 CACGCCATATTCACAAATAAAGAAGTCCAGCTTGACAGATTTTGTAGCCACTCAGTAGG	2400
Qy	2401	AGTCTGAATGCTGATGTGCCAGAACTTATAACTCCATTAGTTTCATTGGGCCACATT	2460
Db	2401	 AGTCTGAATGCTGATGTGCCAGAACTTATAACTCCATTAGTTTCATTGGGCCACATT	2460
Qy	2461	TCTATGTTAGCACCAGATCAGTTTGTCTCCCAATGAAATCTGTAGTAGCAAATTTTATT	2520

Db	2461	TCTATGTTAGCACCAGATCAGTTTGCTTCCCAATGAAATCTGTAGTAGCAAAATTTTATT	2520
Qy	2521	GTGAAAGATCTGCTAATGAATGACAGGTCAACAGGTGAAAAGAATGGAAAACCTGTGGTCT	2580
Db	2521	GTGAAAGATCTGCTAATGAATGACAGGTCAACAGGTGAAAAGAATGGAAAACCTGTGGTCT	2580
Qy	2581	CCAGATGAAGAGGTTTCCCCTGAAGTACTAGCAAAGGTACAGGCAATTAACTTCTGGTA	2640
Db	2581	CCAGATGAAGAGGTTTCCCCTGAAGTACTAGCAAAGGTACAGGCAATTAACTTCTGGTA	2640
Qy	2641	AGGTGGCTGTTGGGTATGAAAAACAACAGTCTAAATCTGCCAATTCAACCTTCGGTTA	2700
Db	2641	AGGTGGCTGTTGGGTATGAAAAACAACAGTCTAAATCTGCCAATTCAACCTTCGGTTA	2700
Qy	2701	TTATCAGCGATGTTGGTTAGTGAGGGTGACCTGACAGAGCAAAAGAGGATCAGTAAATCT	2760
Db	2701	TTATCAGCGATGTTGGTTAGTGAGGGTGACCTGACAGAGCAAAAGAGGATCAGTAAATCT	2760
Qy	2761	GATATGTCTCGCTTGCGATTAGCTGCTGGTAGTGCCATAATGAAGCTTGCTCAGGAACCT	2820
Db	2761	GATATGTCTCGCTTGCGATTAGCTGCTGGTAGTGCCATAATGAAGCTTGCTCAGGAACCT	2820
Qy	2821	TGTTACCATGAAATTATTACCCAGAACAGTTTCAGCTCTGTGCACTTGTTATTAATGAT	2880
Db	2821	TGTTACCATGAAATTATTACCCAGAACAGTTTCAGCTCTGTGCACTTGTTATTAATGAT	2880
Qy	2881	GAGTGTTACCAAGTAAGGCAGATAATTTGCTCAGAAGCTGCATAAGGCACTTGTAAGTTA	2940
Db	2881	GAGTGTTACCAAGTAAGGCAGATAATTTGCTCAGAAGCTGCATAAGGCACTTGTAAGTTA	2940
Qy	2941	CTGCTCCCATTTGGAGTATATGGCGATCTTTGCCTTGTTGTGCCAAAGATCCTGTGAAGGAG	3000
Db	2941	CTGCTCCCATTTGGAGTATATGGCGATCTTTGCCTTGTTGTGCCAAAGATCCTGTGAAGGAG	3000
Qy	3001	AGAAGAGCACACGCACGACAATGTTTACTGAAAAATATCAGTATACGCAGGGAATACATT	3060
Db	3001	AGAAGAGCACACGCACGACAATGTTTACTGAAAAATATCAGTATACGCAGGGAATACATT	3060
Qy	3061	AAGCAGAATCCTATGGCTACTGAGAAATTATTATCACTGTTGCCTGAATATGTAGTTCCA	3120
Db	3061	AAGCAGAATCCTATGGCTACTGAGAAATTATTATCACTGTTGCCTGAATATGTAGTTCCA	3120
Qy	3121	TACATGATTACCTGCTAGCCCATGATCCAGATTTTACAAGATCACAAGATGTTGATCAG	3180
Db	3121	TACATGATTACCTGCTAGCCCATGATCCAGATTTTACAAGATCACAAGATGTTGATCAG	3180
Qy	3181	CTTCGTGATATCAAAGAGTGCCTATGGTTCATGCTTGAAGTTTAAATGACAAAGAATGAA	3240
Db	3181	CTTCGTGATATCAAAGAGTGCCTATGGTTCATGCTTGAAGTTTAAATGACAAAGAATGAA	3240
Qy	3241	AACAATAGCCATGCCTTTATGAAGAAGATGGCAGAGAACATCAAGTTAACCAGAGATGCC	3300
Db	3241	AACAATAGCCATGCCTTTATGAAGAAGATGGCAGAGAACATCAAGTTAACCAGAGATGCC	3300
Qy	3301	CAGTCTCCAGATGAATCCAAGACAAATGAAAACTGTATACAGTATGTGATGTGGCTCTC	3360
Db	3301	CAGTCTCCAGATGAATCCAAGACAAATGAAAACTGTATACAGTATGTGATGTGGCTCTC	3360
Qy	3361	TGTGTTATAAATAGTAAAAGTGCTTTGTGCAATGCAGATTCACCAAAGGACCCAGTCCTC	3420
Db	3361	TGTGTTATAAATAGTAAAAGTGCTTTGTGCAATGCAGATTCACCAAAGGACCCAGTCCTC	3420

Qy	3421	CCAATGAAATTTTACACAACCTGAAAAGGACTTCTGTAACGATAAGAGTTATATTTCA	3480
Db	3421	CCAATGAAATTTTACACAACCTGAAAAGGACTTCTGTAACGATAAGAGTTATATTTCA	3480
Qy	3481	GAAGAGACAAGAGTACTTCTGTTAACAGGAAAGCCAAAGCCTGCTGGAGTACTAGGTGCA	3540
Db	3481	GAAGAGACAAGAGTACTTCTGTTAACAGGAAAGCCAAAGCCTGCTGGAGTACTAGGTGCA	3540
Qy	3541	GTAAATAAGCCTTTATCAGCAACGGGAAGGAAACCTATGTTAGAAGCACTGGCACTGAG	3600
Db	3541	GTAAATAAGCCTTTATCAGCAACGGGAAGGAAACCTATGTTAGAAGCACTGGCACTGAG	3600
Qy	3601	ACTGGAAGCAATATTAATGTAAATTCAGAGCTGAACCTTCAACCGGAAATCGATCAAGG	3660
Db	3601	ACTGGAAGCAATATTAATGTAAATTCAGAGCTGAACCTTCAACCGGAAATCGATCAAGG	3660
Qy	3661	GAACAGAGTTCAGAGGCAGCAGAACTGGAGTTAGTGAAAATGAAGAGAACCTGTGAGG	3720
Db	3661	GAACAGAGTTCAGAGGCAGCAGAACTGGAGTTAGTGAAAATGAAGAGAACCTGTGAGG	3720
Qy	3721	ATTATTTCAGTCAACCTGTAAAGAATATTGACCCAGTAAAGAATAAGGAAATTAATTCT	3780
Db	3721	ATTATTTCAGTCAACCTGTAAAGAATATTGACCCAGTAAAGAATAAGGAAATTAATTCT	3780
Qy	3781	GATCAGGCTACCCAGGGCAACATCAGCAGTGACCGAGGAAAGAAAAGAACAGTAACAGCA	3840
Db	3781	GATCAGGCTACCCAGGGCAACATCAGCAGTGACCGAGGAAAGAAAAGAACAGTAACAGCA	3840
Qy	3841	GCTGGTGAGAGAAATATCCAACAAAAACAGATGAGAAAGTAGATGAATCGGGACCTCCC	3900
Db	3841	GCTGGTGAGAGAAATATCCAACAAAAACAGATGAGAAAGTAGATGAATCGGGACCTCCC	3900
Qy	3901	GCCCCCTCCAAACCCAGGAGAGGACGTCGACCCAAGTCTGAATCTCAGGGCAATGCTACC	3960
Db	3901	GCCCCCTCCAAACCCAGGAGAGGACGTCGACCCAAGTCTGAATCTCAGGGCAATGCTACC	3960
Qy	3961	AAAAATGATGATCTAAATAAACCTATTAACAAGGGAAGGAAGAGAGCTGCAGTGGGTCAG	4020
Db	3961	AAAAATGATGATCTAAATAAACCTATTAACAAGGGAAGGAAGAGAGCTGCAGTGGGTCAG	4020
Qy	4021	GAGAGCCCTGGGGGTTTGAAGCAGGTAATGCCAAAGCACCCAACTGCAAGATTTAGCC	4080
Db	4021	GAGAGCCCTGGGGGTTTGAAGCAGGTAATGCCAAAGCACCCAACTGCAAGATTTAGCC	4080
Qy	4081	AAAAAGGCAGCACCAGCAGAAAGACAAATTGACTTACAAAGGTAAAAATGCATTTGCAAA	4140
Db	4081	AAAAAGGCAGCACCAGCAGAAAGACAAATTGACTTACAAAGGTAAAAATGCATTTGCAAA	4140
Qy	4141	GGGAGAAAAATGAAGGCCAAACAGAAGCAGGCTCCAGCTTCTGCAAAAACTTGGATTACA	4200
Db	4141	GGGAGAAAAATGAAGGCCAAACAGAAGCAGGCTCCAGCTTCTGCAAAAACTTGGATTACA	4200
Qy	4201	AATGTCCCTGAACAGAAAAATGAAGCTCACTTCAGAACACACACTCTCTGCCTTGAAAACT	4260
Db	4201	AATGTCCCTGAACAGAAAAATGAAGCTCACTTCAGAACACACACTCTCTGCCTTGAAAACT	4260
Qy	4261	AAAGAGACTATTACTTCCTTTTCATGACCACAAGTCCTCTGATGGAAATGTACAGCAG	4320
Db	4261	AAAGAGACTATTACTTCCTTTTCATGACCACAAGTCCTCTGATGGAAATGTACAGCAG	4320



Qy	4321	AAACTCTTGAGAGAGAGGCTAAAAGCAACTCTGTTCTCCCCCTTCCCCTAGACTTTTCTT	4380
Db	4321	AAACTCTTGAGAGAGAGGCTAAAAGCAACTCTGTTCTCCCCCTTCCCCTAGACTTTTCTT	4380
Qy	4381	ACGAAAAGTCAATAATTAAGCAAATTGCTTAACACTTGGTTCCAGTTCCTGCCTATCTGG	4440
Db	4381	ACGAAAAGTCAATAATTAAGCAAATTGCTTAACACTTGGTTCCAGTTCCTGCCTATCTGG	4440
Qy	4441	AGTTTAAATGCGTAATACACCATTAAATTTCCACGCTGCAGTTTTTATTTTAAAGAAAGTA	4500
Db	4441	AGTTTAAATGCGTAATACACCATTAAATTTCCACGCTGCAGTTTTTATTTTAAAGAAAGTA	4500
Qy	4501	ACAAGATGTCTTTACACTGACACTGAAAATTCATCCATTTTAGAGCCAGGAATCCCATG	4560
Db	4501	ACAAGATGTCTTTACACTGACACTGAAAATTCATCCATTTTAGAGCCAGGAATCCCATG	4560
Qy	4561	TTACACAGGAAAAATAGAAAGTCTACTGAATTAATTTTTTAAAGAAAAGAGATCAGATT	4620
Db	4561	TTACACAGGAAAAATAGAAAGTCTACTGAATTAATTTTTTAAAGAAAAGAGATCAGATT	4620
Qy	4621	AAATATTTCTTTGTTTTCTTTTGGAACTTTTATGTATAATTCTTTCTGCCTGCCTAC	4680
Db	4621	AAATATTTCTTTGTTTTCTTTTGGAACTTTTATGTATAATTCTTTCTGCCTGCCTAC	4680
Qy	4681	TTTTCTGCAAAAATGAGATGTACAGATTTTCGGTTCCTGCTATGAAAAGTGATGTGGTAG	4740
Db	4681	TTTTCTGCAAAAATGAGATGTACAGATTTTCGGTTCCTGCTATGAAAAGTGATGTGGTAG	4740
Qy	4741	CAATTTTATAAATGTTGCTTTCTGATTTTTATCAGAGTGAGAAAATTAAAATTATTGATT	4800
Db	4741	CAATTTTATAAATGTTGCTTTCTGATTTTTATCAGAGTGAGAAAATTAAAATTATTGATT	4800
Qy	4801	TGCAAGTAGTAAACAGTTCATATTTTGATTTCCCCTCATTTTAGTTTAATATAATTTGCA	4860
Db	4801	TGCAAGTAGTAAACAGTTCATATTTTGATTTCCCCTCATTTTAGTTTAATATAATTTGCA	4860
Qy	4861	ATAAATGTACATATTGTTGTTTGTTCATAAAGCATATCACTTTAAATGGTTTTTACTC	4920
Db	4861	ATAAATGTACATATTGTTGTTTGTTCATAAAGCATATCACTTTAAATGGTTTTTACTC	4920
Qy	4921	CTGTGATTATGTTGGAATATTTGGAATTTTAAAGGAGTAAAGACTGTCCAGCATTGGTT	4980
Db	4921	CTGTGATTATGTTGGAATATTTGGAATTTTAAAGGAGTAAAGACTGTCCAGCATTGGTT	4980
Qy	4981	TTATAATGTTTGTCAACAGATTTTATTAATGTAAAAAATCAATTTTAAAAAATAGTTG	5040
Db	4981	TTATAATGTTTGTCAACAGATTTTATTAATGTAAAAAATCAATTTTAAAAAATAGTTG	5040
Qy	5041	GACTTTGGCAGCTTTTAAGGAAAGTTGGAGGTGTTTAGGATTGCTATCAATTTTCAGCA	5100
Db	5041	GACTTTGGCAGCTTTTAAGGAAAGTTGGAGGTGTTTAGGATTGCTATCAATTTTCAGCA	5100
Qy	5101	TTGTGCTATTTGGAAATAAGTGTTTTGCTTTTGTCTGATGGTCTGGGCTCATTTTTATGT	5160
Db	5101	TTGTGCTATTTGGAAATAAGTGTTTTGCTTTTGTCTGATGGTCTGGGCTCATTTTTATGT	5160
Qy	5161	TTATTTTAGAAAACTGTTGCATCAATATATTATGTTTCTTGGCATTGTTTCAGCATAGGTA	5220
Db	5161	TTATTTTAGAAAACTGTTGCATCAATATATTATGTTTCTTGGCATTGTTTCAGCATAGGTA	5220
Qy	5221	ATGTGTGCACTTTATGTGTACACATAATCATATTTAAGTTTTTGCATAAAATAAATGCT	5280

Db	5221	 ATGTGTGCACTTTATGTGTACACATAATCATATTTAAGTTTTTGCATAAAATAAATGCT	5280
Qy	5281	TCTAGATGTCATGGCAGTCTTTTTAATCTTTTATCATATGCTTTCTTGTGAATTTTTTC	5340
Db	5281	 TCTAGATGTCATGGCAGTCTTTTTAATCTTTTATCATATGCTTTCTTGTGAATTTTTTC	5340
Qy	5341	ATGTTAAAGAGCTAAAGTCATAACATGATTACAGTCAACTCTCCATTATCTATATAAAAT	5400
Db	5341	 ATGTTAAAGAGCTAAAGTCATAACATGATTACAGTCAACTCTCCATTATCTATATAAAAT	5400
Qy	5401	AGTGACTAAGCCTCAGGTTTTTAATTTTGTGATAACAAAATAACGAAGGCATGTAAGACC	5460
Db	5401	 AGTGACTAAGCCTCAGGTTTTTAATTTTGTGATAACAAAATAACGAAGGCATGTAAGACC	5460
Qy	5461	TGATTCTGGAGGAACATGAAATTTGTCTTTTCTCATGTCCAGAGTTCTATCCTGCCCCCA	5520
Db	5461	 TGATTCTGGAGGAACATGAAATTTGTCTTTTCTCATGTCCAGAGTTCTATCCTGCCCCCA	5520
Qy	5521	CTGTCCACTGTAGGGTCATCCGCAAAGCCCTAGCAGAAATGTGCTCACTCCATTTCTTAC	5580
Db	5521	 CTGTCCACTGTAGGGTCATCCGCAAAGCCCTAGCAGAAATGTGCTCACTCCATTTCTTAC	5580
Qy	5581	ACGTTTCTAGCATGGGTGAGAGGAAACAACATTTGTGTTATAACTTCGTCTTGATAGGCT	5640
Db	5581	 ACGTTTCTAGCATGGGTGAGAGGAAACAACATTTGTGTTATAACTTCGTCTTGATAGGCT	5640
Qy	5641	GTAGTGACATGGGATGTAAACAAACAAGTGATCAAAGGTGGATGATTCTGTTAGAGT	5700
Db	5641	 GTAGTGACATGGGATGTAAACAAACAAGTGATCAAAGGTGGATGATTCTGTTAGAGT	5700
Qy	5701	GAAGTTTGAGAGTAAATGTCACTTACGTTTCTCATAGATGATCAAGAGTTGGCTGTGTAT	5760
Db	5701	 GAAGTTTGAGAGTAAATGTCACTTACGTTTCTCATAGATGATCAAGAGTTGGCTGTGTAT	5760
Qy	5761	TGACTGAAAGATGGGTAATTATTTTAAATATGCATTACACACATTTAGGTATCAGAAGA	5820
Db	5761	 TGACTGAAAGATGGGTAATTATTTTAAATATGCATTACACACATTTAGGTATCAGAAGA	5820
Qy	5821	TGCTTAGGGAACAATGGATACCAATGATAGAAAAATGATACCTTTACAGGGGCAGAAAAAT	5880
Db	5821	 TGCTTAGGGAACAATGGATACCAATGATAGAAAAATGATACCTTTACAGGGGCAGAAAAAT	5880
Qy	5881	CCCCACTCTTCCTTATTGCCTCTTCAGAACCCTTTAGAAAGTATAAAATATTGCCTCCAA	5940
Db	5881	 CCCCACTCTTCCTTATTGCCTCTTCAGAACCCTTTAGAAAGTATAAAATATTGCCTCCAA	5940
Qy	5941	CATGCTGAAAAAGAGTATCTATGCATAAGTATCAGAGAAGTCCCTCAAGCAATCAGTAGG	6000
Db	5941	 CATGCTGAAAAAGAGTATCTATGCATAAGTATCAGAGAAGTCCCTCAAGCAATCAGTAGG	6000
Qy	6001	TGTGTTCTATTTAGAGAGAGTTTAAAGTTCTCTTAGCATCAGACAACTTGATTCTTAAGG	6060
Db	6001	 TGTGTTCTATTTAGAGAGAGTTTAAAGTTCTCTTAGCATCAGACAACTTGATTCTTAAGG	6060
Qy	6061	TTTCCAGTGTGTCACCAACAAAAAGTGCATTGATAGGGACCTTTGTCTCTTCCTCCCTTT	6120
Db	6061	 TTTCCAGTGTGTCACCAACAAAAAGTGCATTGATAGGGACCTTTGTCTCTTCCTCCCTTT	6120
Qy	6121	GATTAATTGCCCGGCATCACAGTTTACTAGATTACCAAGTGTTACATCATATTAAATAAA	6180

```
Db      6121 GATTAATTGCCCGGCATCACAGTTTACTAGATTACCAAGTGTTACATCATATTAAATAAA 6180
Qy      6181 ATGTAGCAGAACCATCTGCATCAATATATTCTGTTTAGATTTTTCAGGAGAGAAGTTA 6240
        |||
Db      6181 ATGTAGCAGAACCATCTGCATCAATATATTCTGTTTAGATTTTTCAGGAGAGAAGTTA 6240
Qy      6241 AAAGGATTTGCTCCTTGTATGATGTAAGTGGCCCAACCCCAATTTTGTAACATGATGCAAG 6300
        |||
Db      6241 AAAGGATTTGCTCCTTGTATGATGTAAGTGGCCCAACCCCAATTTTGTAACATGATGCAAG 6300
Qy      6301 TGTCTGGCACTAAGGGAAGCAAGAGTAGGGTTGTGGAAAGACCAAGCTGATGGGGAGGGA 6360
        |||
Db      6301 TGTCTGGCACTAAGGGAAGCAAGAGTAGGGTTGTGGAAAGACCAAGCTGATGGGGAGGGA 6360
Qy      6361 CTTGTTTACGGGAATTTTTTTTAGTTTTCTTTTCAAAGGAAAACATTAAAATCCCTTAGG 6420
        |||
Db      6361 CTTGTTTACGGGAATTTTTTTTAGTTTTCTTTTCAAAGGAAAACATTAAAATCCCTTAGG 6420
Qy      6421 AATTTGGTATTACATCTCAGAGAACTACAACACAAAAGTGCAGACTTATATTTGAGAAT 6480
        |||
Db      6421 AATTTGGTATTACATCTCAGAGAACTACAACACAAAAGTGCAGACTTATATTTGAGAAT 6480
Qy      6481 TAATGTTAACCCCTTTGTGTCTAGTTTGAAGCTTCTTGTATTTGTCTAAACAACAAGCCA 6540
        |||
Db      6481 TAATGTTAACCCCTTTGTGTCTAGTTTGAAGCTTCTTGTATTTGTCTAAACAACAAGCCA 6540
Qy      6541 GAATTTTGTATCTCCTTTGATAAAAAGTGTGTATAATGTAAAGTAGTTTTGCATATTCTT 6600
        |||
Db      6541 GAATTTTGTATCTCCTTTGATAAAAAGTGTGTATAATGTAAAGTAGTTTTGCATATTCTT 6600
Qy      6601 GTGCTGCACATGGGCTGAATTTTTAAATTTTTTTTAAAAACTTGAAGCAGAACCTTGTA 6660
        |||
Db      6601 GTGCTGCACATGGGCTGAATTTTTAAATTTTTTTTAAAAACTTGAAGCAGAACCTTGTA 6660
Qy      6661 TTTGTGTAAATGACAAGTGTAATAATCCTACCATAAAATGCTAAAAATATGCACTGTTTCA 6720
        |||
Db      6661 TTTGTGTAAATGACAAGTGTAATAATCCTACCATAAAATGCTAAAAATATGCACTGTTTCA 6720
Qy      6721 AATAAAACCAAGAAATGCAGCATT 6744
        |||
Db      6721 AATAAAACCAAGAAATGCAGCATT 6744
```

## RESULT 2

ADX06262

ID ADX06262 standard; DNA; 6744 BP.

XX

AC ADX06262;

XX

DT 21-APR-2005 (first entry)

XX

DE Cyclin-dependent kinase modulation biomarker DNA SEQ ID NO 827.

XX

KW cytostatic; cyclin-dependent kinase; cdk; biomarker; gene; ds.

XX

OS Homo sapiens.

XX

PN WO2005012875-A2.

XX

PD 10-FEB-2005.

XX

PF 29-JUL-2004; 2004WO-US024424.  
XX  
PR 29-JUL-2003; 2003US-0490890P.  
XX  
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
PI Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;  
XX  
DR WPI; 2005-163068/17.  
DR P-PSDB; ADX06263.  
XX  
PT Biomarkers useful for predicting or determining the response of a mammal  
PT to a cancer treatment comprising administration of a modulator of cyclin-  
PT dependent kinase activity.  
XX  
PS Claim 5; SEQ ID NO 827; 141pp; English.  
XX  
CC This invention describes a novel method of predicting or determining  
CC whether a mammal will respond or is responding to an anti-cancer agent  
CC that modulates cyclin-dependent kinase (cdk) activity. The method  
CC comprises measuring the level of one or more biomarkers selected from  
CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID  
CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the  
CC invention is utilized in a kit for determining or predicting whether  
CC patient would be susceptible or resistant to treatment by an agent  
CC modulating cdk activity. The invention also describes a method for  
CC utilizing individualized genetic profiles for treating diseases and  
CC disorders based on patient's response and molecular level, specialized  
CC microarrays comprising the biomarkers described, antibodies directed  
CC against the biomarkers and a cell culture model to identify biomarkers.  
CC The cdk modulator is preferably N-5-[[5-(1,1-Dimethylethyl)-2-  
CC oxazolyl]methyl]thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-  
CC tartaric acid salt. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences. This  
CC sequence encodes a biomarker used in the method of the invention.  
XX  
SQ Sequence 6744 BP; 2156 A; 1243 C; 1400 G; 1945 T; 0 U; 0 Other;

Query Match 100.0%; Score 6744; DB 14; Length 6744;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 6744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 GGACCTCGCAGGCCAAGAATTCGGCACGAGGGGCGCCGGCTCCCGGGGCACGGACGGCCG 60
      |||
Db      1 GGACCTCGCAGGCCAAGAATTCGGCACGAGGGGCGCCGGCTCCCGGGGCACGGACGGCCG 60

Qy     61 GGC GCG CGC CTCTGCGAGGGGCGTCCGGGTCCGAGTCGGCGGTCCGGGCCGGCGCGAGGT 120
      |||
Db     61 GGC GCG CGC CTCTGCGAGGGGCGTCCGGGTCCGAGTCGGCGGTCCGGGCCGGCGCGAGGT 120

Qy    121 GCGTGC GGGCGGGCCGCGGGGGTCCCGACGGACACAAGCGCACACACTCCCGGAAGATC 180
      |||
Db    121 GCGTGC GGGCGGGCCGCGGGGGTCCCGACGGACACAAGCGCACACACTCCCGGAAGATC 180

Qy    181 GCTTACCCTCCGGGGGTAAAAGAGATCACCGACAAGATCACCGACGAGATGATCAAA 240
      |||
Db    181 GCTTACCCTCCGGGGGTAAAAGAGATCACCGACAAGATCACCGACGAGATGATCAAA 240

Qy    241 CGCCTGAAGATGGTAGTGAAAACCTTTATGGATATGGATCAGGACTCAGAAGATGAAAAA 300
      |||
```

# SCORE Search Results Details for Application 10679580 and Search Result us-10-679-580a-1.rst.

[Score Home Page](#) [Retrieve Application List](#) [SCORE System Overview](#) [SCORE FAQ](#) [Comments / Suggestions](#)

This page gives you Search Results detail for the Application 10679580 and Search Result us-10-679-580a-1.rst.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2006, 02:05:21 ; Search time 28579 Seconds  
(without alignments)  
13195.702 Million cell updates/sec

Title: US-10-679-580A-1  
Perfect score: 6744  
Sequence: 1 ggacctcgcaggccaagaat.....aaaccaagaaatgcagcatt 6744

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est3:\*  
3: gb\_est4:\*  
4: gb\_est5:\*  
5: gb\_est6:\*  
6: gb\_htc:\*  
7: gb\_est2:\*  
8: gb\_est7:\*  
9: gb\_est8:\*  
10: gb\_est9:\*  
11: gb\_gss1:\*  
12: gb\_gss2:\*  
13: gb\_gss3:\*  
14: gb\_gss4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5070.6	75.2	5134	6	CR749425	CR749425 Homo sapi
2	3628.2	53.8	4019	14	DQ049835	DQ049835 Homo sapi
3	3385.8	50.2	4019	14	DQ049836	DQ049836 Pan trogl
4	932.4	13.8	2899	6	AK041682	AK041682 Mus muscu
5	865.2	12.8	939	3	BQ959509	BQ959509 AGENCOURT
6	852.8	12.6	1182	2	BM454791	BM454791 AGENCOURT
7	850.8	12.6	909	3	BU195798	BU195798 AGENCOURT
8	848.2	12.6	1088	2	BM475383	BM475383 AGENCOURT
9	845.8	12.5	923	3	BQ438312	BQ438312 AGENCOURT
c 10	841.4	12.5	931	10	DV928225	DV928225 LB03012.C
11	833.8	12.4	922	8	CR986733	CR986733 CR986733
12	832.2	12.3	1106	3	BM799673	BM799673 AGENCOURT
13	827.8	12.3	874	3	BU517176	BU517176 AGENCOURT
14	826.2	12.3	859	8	CR999562	CR999562 CR999562
15	819.8	12.2	991	2	BM470743	BM470743 AGENCOURT
16	817.4	12.1	867	3	BU159962	BU159962 AGENCOURT
17	814.6	12.1	1142	2	BM466282	BM466282 AGENCOURT
18	810	12.0	1002	2	BM465904	BM465904 AGENCOURT
19	808.8	12.0	835	9	DA735319	DA735319 DA735319
20	800.4	11.9	856	3	BU166770	BU166770 AGENCOURT
21	791.4	11.7	1056	2	BM477531	BM477531 AGENCOURT
22	782.6	11.6	3313	6	AK045159	AK045159 Mus muscu
23	776.4	11.5	792	8	CN430328	CN430328 170005322
24	770.8	11.4	853	3	BQ440349	BQ440349 AGENCOURT
25	767.4	11.4	1092	2	BM465566	BM465566 AGENCOURT
26	761	11.3	764	9	DA736985	DA736985 DA736985
27	755	11.2	766	8	CN430329	CN430329 170005319
c 28	754.2	11.2	3378	6	AK045737	AK045737 Mus muscu
29	749	11.1	773	1	AL705029	AL705029 DKFZp686B
30	748.8	11.1	979	2	BG758965	BG758965 602713079
31	747.8	11.1	786	2	BG208101	BG208101 RST27591
32	745	11.0	864	8	CV884014	CV884014 UI-M-HZ0-
c 33	743.8	11.0	780	3	BU630464	BU630464 UI-H-FL0-
34	741	11.0	796	1	AU137771	AU137771 AU137771
35	731.6	10.8	1047	8	CO647378	CO647378 ILLUMIGEN
c 36	729.6	10.8	766	3	BU631648	BU631648 UI-H-FL0-
37	729	10.8	761	8	CN430334	CN430334 170006001
38	728	10.8	825	2	BI254793	BI254793 602974029
39	726.8	10.8	785	1	AU137328	AU137328 AU137328
c 40	720.8	10.7	745	3	BU631677	BU631677 UI-H-FL0-
41	720.4	10.7	752	8	CN430332	CN430332 170004248
42	719	10.7	719	9	CX868031	CX868031 HESC4_25_
43	718	10.6	878	3	BM801605	BM801605 AGENCOURT
44	715.2	10.6	738	1	AL045878	AL045878 DKFZp434J
45	712.4	10.6	823	3	BU703952	BU703952 UI-M-FO0-

## ALIGNMENTS

## RESULT 1

CR749425

LOCUS CR749425 5134 bp mRNA linear HTC 17-APR-2005

DEFINITION Homo sapiens mRNA; cDNA DKFZp686B19246 (from clone DKFZp686B19246).

ACCESSION CR749425

VERSION CR749425.1 GI:51476560

KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1 (bases 1 to 5134)

AUTHORS Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,  
Fobo, G., Han, M. and Wiemann, S.

CONSRTM The German cDNA Consortium

TITLE Direct Submission

JOURNAL Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764  
Neuherberg, GERMANY

COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by BMFZ (Biomedical Research Center at the  
Heinrich-Heine-University, Duesseldorf/Germany) within the cDNA  
sequencing consortium of the German Genome Project. This clone  
(DKFZp686B19246) is available at the RZPD Deutsches  
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.  
Please contact RZPD for ordering:  
<http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686B19246>  
Further information about the clone and the sequencing project is  
available at <http://mips.gsf.de/projects/cdna/>.

FEATURES Location/Qualifiers

source 1. .5134  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp686B19246"  
/tissue\_type="salivary gland"  
/clone\_lib="686 (synonym: hlcc3). Vector pSport1\_Sfi; host  
DH10B; sites SfiIA + SfiIB"  
/dev\_stage="adult"  
/note="hypothetical protein"

gene 1. .5134  
/gene="DKFZp686B19246"

CDS 30. .3923  
/gene="DKFZp686B19246"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAH18263.1"  
/db\_xref="GI:51476561"  
/db\_xref="GOA:Q68DF7"  
/db\_xref="InterPro:IPR000357"  
/db\_xref="UniProtKB/TrEMBL:Q68DF7"  
/translation="MIKRLKMVVKTFMDMDQDSEDEKQQYLPLALHLASEFFLRNPKNK  
DVRLLVACCLADIFRIYAPEAPYTS HDKLDIFLFI TRQLKGLEDTKSPQFNRYFYLL  
ENLAWVKSYNICFELEDCNEIFIQLFRTLFSVINNSHNKKVQMHLMDLMSIIMEGDG  
VTQELLDSILINLI PAHKNLNKQSF DLAKVLLKRTVQTI EACIANFFNQVLVLGRSSV  
SDLSEHVFDLIQELFAIDPHLLSVMPQLEFKLSNDGEERLAVVRL LAKLFGSKDSD  
LATQNRPLWQCFLGRFNDIHVPVRLESVKFASHCLMNHPDLAKDLTEYLKVRSHDPEE  
AIRHDVIVTI ITAAKRD LALVNDQLLG FVRERTLDKRWVRKEAMMGLAQLYKKYCLH  
GEAGKEAAEKVSWIKDKLLHIYYQNSIDDKLLVEKIFAQYL VPHNLETEERMKCLYYL  
YASLDPNAVKALNEMWK CQNMLRSHVRELLDLH KQPTSEANCSAMFGKLM TIAKNLPD  
PGKAQDFVKKFNQVLG DDEKLRSQLELLISPTC SCKQADICVREIARKLANPKQPTNP  
FLEMVKFLLERIAPVHIDSE AISALVKLMNKSIEGTADDEEEGVS PDTAIRSGLELLK  
VLSFTHPTSFHSAETYESLLQCLRMEDDKVAEAAI QIFRNTGHK IETDLPQIRSTLIP  
ILHQAKRGTPHQAQAVHCIHAIFTNKEVQLAQI FEPLSRSLNADVPEQLITPLVSL  
GHISMLAPDQFASPMKSVVANFIVKDLLMND RSTGEKNGKLWSPDEEVSPEVLAKVQA  
IKLLVRWLLGMKNNQSKSANSTLRLLSAMLVSEGLTEQKRISKSDMSRLRLAAGSAI

MKLAQEPYHEIITPEQFQLCALVINDECYQVRQIFAQKLHKALVKLLLPLEYMAIFA  
 LCAKDPVKERRAHARQC LLKNISIRREYIKQNPMAATEKL LSLPEYVVPYMIHLLAHD  
 PDFTRSQDQDLRDIKECLWFMLEVLMTKNENNSHAFMKMAENIKLTRDAQSPDESK  
 TNEKLYTVCDVALCVINSK SALCNADSPKDPVLP MKFFTQPEKDFCNDKSYISEETRV  
 LLLTGKPKPAGVLGAVNKPLSATGRKPYVRSTGTETG SNINVNSELNPSTGNRSREQS  
 SEAAETGVSENEENPVR IISVTPVKNIDPVKNKEINS DQATQGNISSDRGKKRTVTAA  
 GAENIQQKTDEKVDSESGPPAPSKPRRGRRPKSESQGNATKNDDL NKPINKGRKRAAVG  
 QESPGGLEAGNAKAPKLQDLAKKAAPAERQIDLQR"

## ORIGIN

Query Match 75.2%; Score 5070.6; DB 6; Length 5134;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 5085; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

Qy	203	AGATCACCGACAAGATCACCGACGAGATGATCAAACGCCTGAAGATGGTAGTGAAAA	262
Db	1	AGATCACCGACAAGATCACCGACGAGATGATCAAACGCCTGAAGATGGTAGTGAAAA	60
Qy	263	CCTTTATGGATATGGATCAGGACTCAGAAGATGAAAAACAGCAGTATCTCCCACTAGCCT	322
Db	61	CCTTTATGGATATGGATCAGGACTCAGAAGATGAAAAACAGCAGTATCTCCCACTAGCCT	120
Qy	323	TGCATCTTGCATCTGAATTCTTCCTCAGGAACCCCAATAAAGATGTGCGTCTCCTTG TAG	382
Db	121	TGCATCTTGCATCTGAATTCTTCCTCAGGAACCCCAATAAAGATGTGCGTCTCCTTG TAG	180
Qy	383	CATGTTGTTTGGCTGATATCTTTCGTATCTATGCCCCAGAAGCTCCATATACTTCCCATG	442
Db	181	CATGTTGTTTGGCTGATATCTTTCGTATCTATGCCCCAGAAGCTCCATATACTTCCCATG	240
Qy	443	ATAAACTTAAGGACATATTTTTGT TTTATTACCAGACAATTAAGGTTTGAGGATACAA	502
Db	241	ATAAACTTAAGGACATATTTTTGT TTTATTACCAGACAATTAAGGTTTGAGGATACAA	300
Qy	503	AGAGTCCACAGTTTAATAGATACTTTTATTTATTAGAGAATTTAGCTTGGGT TAAATCAT	562
Db	301	AGAGTCCACAGTTTAATAGATACTTTTATTTATTAGAGAATTTAGCTTGGGT TAAATCAT	360
Qy	563	ATAACATCTGCTTTGAATTGGAAGATTGCAATGAAATTTTATT CAGCTTTT TAGAACTC	622
Db	361	ATAACATCTGCTTTGAATTGGAAGATTGCAATGAAATTTTATT CAGCTTTT TAGAACTC	420
Qy	623	TCTTCTCAGTGATCAACAATAGCCACAATAAGAAGGTACAAATGCACATGCTAGATT TGA	682
Db	421	TCTTCTCAGTGATCAACAATAGCCACAATAAGAAGGTACAAATGCACATGCTAGATT TGA	480
Qy	683	TGAGTTCTATCATCATGGAAGGTGATGGAGTTACTCAAGAATTATTGGGCTCCATTCTTA	742
Db	481	TGAGTTCTATCATCATGGAAGGTGATGGAGTTACTCAAGAATTATTGGGCTCCATTCTTA	540
Qy	743	TTAACCTCATTCTGACATAAGAAGCTTAAATAAACAGTCCTTTGACCTTGCAAAAGTGC	802
Db	541	TTAACCTCATTCTGACATAAGAAGCTTAAATAAACAGTCCTTTGACCTTGCAAAAGTGC	600
Qy	803	TATTGAAAAGAACAGTCCAGACTATTGAGGCATGCATTGCCAATTTTTTCAATCAAGTCC	862
Db	601	TATTGAAAAGAACAGTCCAGACTATTGAGGCATGCATTGCTAATTTTTTCAATCAAGTCC	660
Qy	863	TGGTGCTGGGAAGATCATCAGTAAGTGATTGTGTCAGAACATGTATTTGATCTGATT CAGG	922
Db	661	TGGTGCTGGGAAGATCATCAGTAAGTGATTGTGTCAGAACATGTATTTGATCTGATT CAGG	720



SCORE Search Results Details for Application 10679580 and Search Result us-10-679-580a-2.rag.

[Score Home Page](#) [Retrieve Application List](#) [SCORE System Overview](#) [SCORE FAQ](#) [Comments / Suggestions](#)

This page gives you Search Results detail for the Application 10679580 and Search Result us-10-679-580a-2.rag.

[start](#)

Go Back to previous page

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2006, 21:02:13 ; Search time 203 Seconds  
(without alignments)  
2921.230 Million cell updates/sec

Title: US-10-679-580A-2  
Perfect score: 6650  
Sequence: 1 MIKRLKMVVKTFFMDMDQDSE.....KLQDLAKKAAPAERQIDLQR 1297

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_8:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*  
9: geneseqp2005s:\*  
10: geneseqp2006s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
%					
Result	Query				
No.	Score	Match	Length	DB ID	Description

1	6650	100.0	1297	9	ADX06263	Adx06263	Cyclin-de
2	6643	99.9	1337	6	ABR58297	Abr58297	BCU0720 p
3	6618	99.5	1297	6	ABG72803	Abg72803	Human tum
4	4711	70.8	919	5	ABP64685	Abp64685	Human pro
5	4668.5	70.2	1391	3	AAV94702	Aay94702	Human and
6	4363	65.6	854	3	AAB01382	Aab01382	Neuron-as
7	3494	52.5	687	6	ABR58332	Abr58332	XM_047529
8	2842	42.7	600	7	ADM06198	Adm06198	Human pro
9	2842	42.7	600	9	AEC89128	Aec89128	Human cDN
10	2439.5	36.7	846	7	ADJ69898	Adj69898	Human hea
11	1911	28.7	1218	4	ABB67869	Abb67869	Drosophil
12	1861	28.0	363	3	AAB58462	Aab58462	Lung canc
13	1715	25.8	333	4	AAB94946	Aab94946	Human pro
14	1440.5	21.7	588	8	ADR10188	Adr10188	Human pro
15	1187	17.8	233	4	AAB95449	Aab95449	Human pro
16	910	13.7	182	5	ABB84560	Abb84560	Human euc
17	862	13.0	165	4	AAU01759	Aau01759	Human sec
18	759	11.4	161	4	AAU01760	Aau01760	Human sec
19	387	5.8	147	5	ABB89260	Abb89260	Human pol
20	377	5.7	92	3	AAG02811	Aag02811	Human sec
21	314	4.7	101	3	AAV65247	Aay65247	Human 5'
22	314	4.7	101	8	ADU72811	Adu72811	Signal pe
23	314	4.7	101	9	ADZ73802	Adz73802	Human inc
24	217.5	3.3	1743	2	AAW98879	Aaw98879	H. pylori
25	214.5	3.2	1875	6	ABR53560	Abr53560	Protein s
26	214.5	3.2	1875	7	ADK64380	Adk64380	Disease t
27	214.5	3.2	1875	8	ADS43855	Ads43855	Bacterial
28	213	3.2	761	8	ADX79770	Adx79770	Plant ful
29	204.5	3.1	1790	6	ABR53116	Abr53116	Protein s
30	204.5	3.1	1790	7	ADK63040	Adk63040	Disease t
31	204.5	3.1	1790	8	ADN18933	Adn18933	Bacterial
32	204.5	3.1	1790	9	ADY85285	Ady85285	Yeast hom
33	204	3.1	759	7	ABM88120	Abm88120	Rice abio
34	199.5	3.0	960	4	AAB94466	Aab94466	Human pro
35	199.5	3.0	1015	4	AAG64040	Aag64040	Human DRF
36	199.5	3.0	1015	9	ADY20095	Ady20095	PRO polyp
37	199.5	3.0	1015	9	ADY17460	Ady17460	PRO polyp
38	199.5	3.0	1160	4	ABB59418	Abb59418	Drosophil
39	199	3.0	2568	9	ADW44253	Adw44253	Human cen
40	199	3.0	2633	4	ABG06505	Abg06505	Novel hum
41	199	3.0	2663	4	AAM39097	Aam39097	Human pol
42	199	3.0	2663	8	ADQ17932	Adq17932	Human sof
43	199	3.0	2663	9	ADX06867	Adx06867	Cyclin-de
44	199	3.0	2688	4	AAM40883	Aam40883	Human pol
45	198.5	3.0	1927	6	ABU30805	Abu30805	Protein e

## ALIGNMENTS

## RESULT 1

ADX06263

ID ADX06263 standard; protein; 1297 AA.

XX

AC ADX06263;

XX

DT 21-APR-2005 (first entry)

XX

DE Cyclin-dependent kinase modulation biomarker SEQ ID NO 828.

XX

Query Match 100.0%; Score 6650; DB 9; Length 1297;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

Db	121	NEIFIQLFRTLFSVINNSHNKKVQMHMLDLMSSIIMEGDGVTQELLGSILINLI PAHKNL	180
Qy	181	NKQSFDLAKVLLKRTVQTI EACIANFFNQVLVLGRSSVSDLSEHVFDLIQELFAIDPHLL	240
Db	181	NKQSFDLAKVLLKRTVQTI EACIANFFNQVLVLGRSSVSDLSEHVFDLIQELFAIDPHLL	240
Qy	241	LSVMPQLEFKLKSNDGEERLAVVRL LAKLFGSKDSLATQNRPLWQCF LGRFNDIHVPVR	300
Db	241	LSVMPQLEFKLKSNDGEERLAVVRL LAKLFGSKDSLATQNRPLWQCF LGRFNDIHVPVR	300
Qy	301	LESVKFASHCLMNHPDLAKDLTEYLKVRSHDPEEAI RHDVIVTI ITAAKRDALVNDQLL	360
Db	301	LESVKFASHCLMNHPDLAKDLTEYLKVRSHDPEEAI RHDVIVTI ITAAKRDALVNDQLL	360
Qy	361	GFVRERTLDKRWRVRKEAMMGLAQLYKKYCLHGEAGKEAAEKVSWIKDKLLHIYYQNSID	420
Db	361	GFVRERTLDKRWRVRKEAMMGLAQLYKKYCLHGEAGKEAAEKVSWIKDKLLHIYYQNSID	420
Qy	421	DKLLVEKIFAQYLVPHNLETEERMKCLYYLYASLDPNAV KALNEMWKCQNMLRSHVRELL	480
Db	421	DKLLVEKIFAQYLVPHNLETEERMKCLYYLYASLDPNAV KALNEMWKCQNMLRSHVRELL	480
Qy	481	DLHKQPTSEANCSAMFGKLM TIAKNLDPGKAQDFVKKFNQVLGDDEKLRSQLELLISPT	540
Db	481	DLHKQPTSEANCSAMFGKLM TIAKNLDPGKAQDFVKKFNQVLGDDEKLRSQLELLISPT	540
Qy	541	CSCQADICVREIARKLANPKQPTNPFLEVMKFLLERIAPVHIDSEAI SALVKLMNKSIE	600
Db	541	CSCQADICVREIARKLANPKQPTNPFLEVMKFLLERIAPVHIDSEAI SALVKLMNKSIE	600
Qy	601	GTADDEEEGVSPDTAIRSGLELLKVLSFTHPTS FHS AETYESLLQCLRMEDDKVAEAAIQ	660
Db	601	GTADDEEEGVSPDTAIRSGLELLKVLSFTHPTS FHS AETYESLLQCLRMEDDKVAEAAIQ	660
Qy	661	IFRNTGHK IETDLPQIRSTLIPILHQAKRGTPHQAKQAVHCIHAI FTNKEVQLAQIFEP	720
Db	661	IFRNTGHK IETDLPQIRSTLIPILHQAKRGTPHQAKQAVHCIHAI FTNKEVQLAQIFEP	720
Qy	721	LSRSLNADVPEQLITPLVSLGHISMLAPDQFASPMKSVVANFIVKD LLMNDRSTGEKNGK	780
Db	721	LSRSLNADVPEQLITPLVSLGHISMLAPDQFASPMKSVVANFIVKD LLMNDRSTGEKNGK	780
Qy	781	LWSPDEEVSPEVLAKVQAIKLLVRWLLGMKNNQSKSANSTLRLLSAMLVSEGDLTEQKRI	840
Db	781	LWSPDEEVSPEVLAKVQAIKLLVRWLLGMKNNQSKSANSTLRLLSAMLVSEGDLTEQKRI	840
Qy	841	SKSDMSRLRLAAGSAIMKLAQEPCYHEIITPEQFQLCALVINDECYQVRQIFAQKLHKAL	900
Db	841	SKSDMSRLRLAAGSAIMKLAQEPCYHEIITPEQFQLCALVINDECYQVRQIFAQKLHKAL	900
Qy	901	VKLLLPLEYMAIFALCAKDPVKERRAHARQCLLKNI SIRREYIKQNP MATEKLLSLLPEY	960
Db	901	VKLLLPLEYMAIFALCAKDPVKERRAHARQCLLKNI SIRREYIKQNP MATEKLLSLLPEY	960
Qy	961	VVPYMIHLLAHDPDFTRSQDQDLRDIKECLWFMLEVLMTKNENNSHAFMKKMAENIKLT	1020
Db	961	VVPYMIHLLAHDPDFTRSQDQDLRDIKECLWFMLEVLMTKNENNSHAFMKKMAENIKLT	1020
Qy	1021	RDAQSPDESKTNEKLYTVCDVALCVINSKSALCNADSPKDPVLP MKFFTQPEKDFCNDKS	1080
Db	1021	RDAQSPDESKTNEKLYTVCDVALCVINSKSALCNADSPKDPVLP MKFFTQPEKDFCNDKS	1080

```

Qy      1081 YISEETRVLLLLTGKPKPAGVLGAVNKPLSATGRKPYVRSTGTETGSNINVNSELNPSTGN 1140
        |||
Db      1081 YISEETRVLLLLTGKPKPAGVLGAVNKPLSATGRKPYVRSTGTETGSNINVNSELNPSTGN 1140

Qy      1141 RSREQSSEAAETGVSENEENPVRIISVTPVKNIDPVKNKEINSDQATQGNISSDRGKKRT 1200
        |||
Db      1141 RSREQSSEAAETGVSENEENPVRIISVTPVKNIDPVKNKEINSDQATQGNISSDRGKKRT 1200

Qy      1201 VTAAGAENIQKTDKVDKESGPPAPSKPRRGRRPKSESQGNATKNDDLKPKPINKGRKRAA 1260
        |||
Db      1201 VTAAGAENIQKTDKVDKESGPPAPSKPRRGRRPKSESQGNATKNDDLKPKPINKGRKRAA 1260

Qy      1261 VGQESPGGLEAGNAKAPKLQDLAKKAAPAERQIDLQR 1297
        |||
Db      1261 VGQESPGGLEAGNAKAPKLQDLAKKAAPAERQIDLQR 1297

```

## RESULT 2

ABR58297

ID ABR58297 standard; protein; 1337 AA.

XX

AC ABR58297;

XX

DT 08-JUL-2003 (first entry)

XX

DE BCU0720 protein #SEQ ID 40.

XX

KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;  
 KW drug discovery; clinical medicine; forensic medicine; chromosome 4p14.

XX

OS Homo sapiens.

XX

PN WO2003029421-A2.

XX

PD 10-APR-2003.

XX

PF 02-OCT-2002; 2002WO-US031287.

XX

PR 03-OCT-2001; 2001US-0326526P.

PR 14-MAY-2002; 2002US-00144194.

XX

PA (ORIG-) ORIGENE TECHNOLOGIES INC.

XX

PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;

XX

DR WPI; 2003-381623/36.

DR N-PSDB; ACC72031.

XX

PT New isolated human differentially-regulated breast cancer polynucleotide  
 PT and polypeptide, useful for diagnosing, staging, prognosticating,  
 PT preventing and/or treating diseases and conditions relating to breast  
 PT cancer.

XX

PS Claim 6; SEQ ID NO 40; 127pp + Sequence Listing; English.

XX

CC The invention relates to isolated polynucleotides which are  
 CC differentially-regulated in breast cancer. The methods and compositions  
 CC of the present invention are useful for detecting, diagnosing, staging,  
 CC monitoring, prognosticating, preventing and/or treating diseases and  
 CC conditions relating to breast cancer, and may be used in gene therapy or

CC antisense therapy. They can also be used in research, drug discovery,  
CC clinical medicine and forensic medicine. Sequences given in records  
CC ABR58278-ABR58323 and ABR58346-ABR58362 represent polypeptides encoded by  
CC polynucleotides of the invention that are differentially-regulated in  
CC breast cancer. NOTE: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX

SQ Sequence 1337 AA;

Query Match 99.9%; Score 6643; DB 6; Length 1337;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1296; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MIKRLKMVVKTFMDMDQDSEDEKQYQLPLALHLASEFFLRNPNDVRLLVACCLADIFRI	60
Db	41	MIKRLKMVVKTFMDMDQDSEDEKQYQLPLALHLASEFFLRNPNDVRLLVACCLADIFRI	100
Qy	61	YAPEAPYTSHDKLKDIFLFITRQLKGLEDTKSPQFNRYFYLLLENLAWVKSYNICFELEDC	120
Db	101	YAPEAPYTSHDKLKDIFLFITRQLKGLEDTKSPQFNRYFYLLLENLAWVKSYNICFELEDC	160
Qy	121	NEIFIQLFRTLFSVINNSHNKKVQMHMLDLMSSI IMEGDGVQTQELLGSILINLI PAHKNL	180
Db	161	NEIFIQLFRTLFSVINNSHNKKVQMHMLDLMSSI IMEGDGVQTQELLDSILINLI PAHKNL	220
Qy	181	NKQSFDLAKVLLKRTVQTI EACIANFFNQVLVLRSSVSDLSEHVFDLIQELFAIDPHLL	240
Db	221	NKQSFDLAKVLLKRTVQTI EACIANFFNQVLVLRSSVSDLSEHVFDLIQELFAIDPHLL	280
Qy	241	LSVMPQLEFKLKSNDGEERLAVVRL LAKLFGSKDSDLATQNRPLWQCFLGRFNDIHPVVR	300
Db	281	LSVMPQLEFKLKSNDGEERLAVVRL LAKLFGSKDSDLATQNRPLWQCFLGRFNDIHPVVR	340
Qy	301	LESVKFASHCLMNHPDLAKDLTEYLKVRSHDPEEAI RHDVIVTI ITAAKRDALVNDQLL	360
Db	341	LESVKFASHCLMNHPDLAKDLTEYLKVRSHDPEEAI RHDVIVTI ITAAKRDALVNDQLL	400
Qy	361	GFVRERTLDRWRVRKEAMMGLAQLYKKYCLHGEAGKEAAEKVSWIKDKLLHIYYQNSID	420
Db	401	GFVRERTLDRWRVRKEAMMGLAQLYKKYCLHGEAGKEAAEKVSWIKDKLLHIYYQNSID	460
Qy	421	DKLLVEKIFAQYLVPHNLETEERMKCLYYLYASLDPNAV KALNEMWKCQNMLRSHVRELL	480
Db	461	DKLLVEKIFAQYLVPHNLETEERMKCLYYLYASLDPNAV KALNEMWKCQNMLRSHVRELL	520
Qy	481	DLHKQPTSEANCSAMFGKLMTIAKNLPDPGKAQDFVKKFNQVLGDDEKLRSQLELLISPT	540
Db	521	DLHKQPTSEANCSAMFGKLMTIAKNLPDPGKAQDFVKKFNQVLGDDEKLRSQLELLISPT	580
Qy	541	CSCQQADICVREIARKLANPKQPTNPFLEMVKFLLERIAPVHIDSEAI SALVKLMNKSIE	600
Db	581	CSCQQADICVREIARKLANPKQPTNPFLEMVKFLLERIAPVHIDSEAI SALVKLMNKSIE	640
Qy	601	GTADDEEEGVSPDTAIRSGLELLKVL SFTHPTS FHS AETYESLLQCLRMEDDKVAEAAIQ	660
Db	641	GTADDEEEGVSPDTAIRSGLELLKVL SFTHPTS FHS AETYESLLQCLRMEDDKVAEAAIQ	700
Qy	661	IFRNTGHKIE TDL PQIRSTLI PILHQAKARGT PHQAKQAVHCIHAI FTNKEVQLAQIFEP	720
Db	701	IFRNTGHKIE TDL PQIRSTLI PILHQAKARGT PHQAKQAVHCIHAI FTNKEVQLAQIFEP	760

```

Qy      721 LSRSLNADVPEQLITPLVSLGHISMLAPDQFASPMKSVVANFIVKDLLMNRSTGEKNGK 780
      |||
Db      761 LSRSLNADVPEQLITPLVSLGHISMLAPDQFASPMKSVVANFIVKDLLMNRSTGEKNGK 820

Qy      781 LWSPDEEVSPPEVLAKVQAIKLLVRWLLGMKNNQSKSANSTLRLLSAMLVSEGDLTEQKRI 840
      |||
Db      821 LWSPDEEVSPPEVLAKVQAIKLLVRWLLGMKNNQSKSANSTLRLLSAMLVSEGDLTEQKRI 880

Qy      841 SKSDMSRLRLAAGSAIMKLAQEPYHEIITPEQFQLCALVINDECYQVRQIFAQKLHKAL 900
      |||
Db      881 SKSDMSRLRLAAGSAIMKLAQEPYHEIITPEQFQLCALVINDECYQVRQIFAQKLHKAL 940

Qy      901 VKLLLPLEYMAIFALCAKDPVKERRAHARQCLLKNISIRREYIKQNPIMATEKLLSLLPEY 960
      |||
Db      941 VKLLLPLEYMAIFALCAKDPVKERRAHARQCLLKNISIRREYIKQNPIMATEKLLSLLPEY 1000

Qy      961 VVPYMIHLLAHDPDFTRSQDQDQLRDIKECLWFMLEVLMTKNENNSHAFMKKMAENIKLT 1020
      |||
Db     1001 VVPYMIHLLAHDPDFTRSQDQDQLRDIKECLWFMLEVLMTKNENNSHAFMKKMAENIKLT 1060

Qy     1021 RDAQSPDESKTNEKLYTVCDVALCVINSKSALCNADSPKDPVLPMKFFTQPEKDFCNDKS 1080
      |||
Db     1061 RDAQSPDESKTNEKLYTVCDVALCVINSKSALCNADSPKDPVLPMKFFTQPEKDFCNDKS 1120

Qy     1081 YISEETRVLLLLTGKPKPAGVLGAVNKPLSATGRKPYVRSTGTETGSNINVNSELNPSTGN 1140
      |||
Db     1121 YISEETRVLLLLTGKPKPAGVLGAVNKPLSATGRKPYVRSTGTETGSNINVNSELNPSTGN 1180

Qy     1141 RSREQSSEAAETGVSENEENPVRIISVTPVKNIDPVKNKEINSDQATQGNISSDRGKKRT 1200
      |||
Db     1181 RSREQSSEAAETGVSENEENPVRIISVTPVKNIDPVKNKEINSDQATQGNISSDRGKKRT 1240

Qy     1201 VTAAGAENIQQKTDEKVDSEGPAPSKPRRGRRPKSESQGNATKNDDLNPINKGRKRAA 1260
      |||
Db     1241 VTAAGAENIQQKTDEKVDSEGPAPSKPRRGRRPKSESQGNATKNDDLNPINKGRKRAA 1300

Qy     1261 VGQESPGGLEAGNAKAPKLQDLAKKAAPAERQIDLQR 1297
      |||
Db     1301 VGQESPGGLEAGNAKAPKLQDLAKKAAPAERQIDLQR 1337

```

## RESULT 3

ABG72803

ID ABG72803 standard; protein; 1297 AA.

XX

AC ABG72803;

XX

DT 24-FEB-2003 (first entry)

XX

DE Human tumour supressor SCC-112.

XX

KW Human; SCC-112; cancer; apoptosis; allodynia; cell proliferation;

KW degenerative disorder; metastasis inhibition; breast cancer; causalgia;

KW kidney cancer; bladder cancer; pancreatic cancer; colon cancer;

KW squamous cell carcinoma; head trauma; spinal cord injury; herpes zoster;

KW global and focal ischaemic and haemorrhagic stroke; epilepsy; neuralgia;

KW hypoxia-induced nerve cell damage; anxiety; diabetes mellitus;

KW cardiac arrest; spinal cord lesion; lung cancer; tumour supressor;

KW neonatal distress; Alzheimer's disease; colon cancer; stomach cancer;

KW multiple sclerosis; phantom limb pain; hyperalgesia; Down's syndrome;

KW Huntington's disease; Parkinson's disease; Korsakoff's syndrome;  
 KW amyotrophic lateral sclerosis; cell survival; uterine cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 2. .137  
 FT /label= RhoGEF domain  
 FT Domain 127. .651  
 FT /label= N-adaptin\_domain  
 FT Region 166. .187  
 FT /label= Leucine\_Zipper  
 FT Domain 249. .350  
 FT /label= SKP1\_domain  
 FT Region 597. .617  
 FT /note= "PEST sequence site"  
 FT Region 858. .865  
 FT /note= "Tyrosine kinase phosphorylation site"  
 FT Region 920. .926  
 FT /note= "Nuclear localisation signature sequence"  
 FT Region 1030. .1036  
 FT /note= "Tyrosine kinase phosphorylation site"  
 FT Region 1143. .1663  
 FT /note= "PEST sequence site"  
 FT Region 1216. .1277  
 FT /note= "PEST sequence site"  
 FT Region 1225. .1231  
 FT /note= "Nuclear localisation signature sequence"  
 FT Region 1227. .1230  
 FT /note= "Nuclear localisation signature sequence"  
 FT Region 1228. .1234  
 FT /note= "Nuclear localisation signature sequence"  
 FT Region 1232. .1235  
 FT /note= "Nuclear localisation signature sequence"  
 FT Region 1251. .1257  
 FT /note= "Nuclear localisation signature sequence"  
 FT Region 1278. .1297  
 FT /label= Epitope\_region  
 FT /note= "Residues 1278 to 1297 are specifically claimed in  
 FT claim 41"  
 XX  
 PN WO200281641-A2.  
 XX  
 PD 17-OCT-2002.  
 XX  
 PF 08-APR-2002; 2002WO-US010850.  
 XX  
 PR 06-APR-2001; 2001US-0281780P.  
 XX  
 PA (GEOU ) UNIV GEORGETOWN.  
 XX  
 PI Kasid UN, Kumar D, Ahmad I;  
 XX  
 DR WPI; 2003-103330/09.  
 DR N-PSDB; ABX14056.  
 XX  
 PT New isolated SSC (undefined) tumor suppressor polypeptides and  
 PT polynucleotides, useful for diagnosing, preventing or treating cancer or  
 PT degenerative disease, e.g. Alzheimer's Disease, Huntington's disease, or  
 PT multiple sclerosis.  
 XX



XX

XX

Qy	481	DLHKQPTSEANCSAMFGKLM	TIAKNLPDPGKAQDFVKKFNQVLGDDEKLRSQLELLISPT	540
Db	481	DLHKQPTSEANCSAMFGKLM	TIAKNLPDPGKAQDFVKKFNQVLGDDEKLRSQLELLISPT	540
Qy	541	CSCQADICVREIARKLANPKQPTNP	FLEMVKFLLERIAPVHIDSEAI	SALVKLMNKSIE 600
Db	541	CSCQADICVREIARKLANPKQPTNP	FLEMVKFLLERIAPVHIDSEAI	SALVKLMNKSIE 600
Qy	601	GTADDEEEGVSPDTAIRSGLELLKVL	SFTHPTSFHSAETYESLLQCLRMEDDKVAEAAIQ	660
Db	601	GTADDEEEGVSPDTAIRSGLELLKVL	SFTHPTSFHSAETYESLLQCLRMEDDKVAEAAIQ	660
Qy	661	IFRNTGHKIE	TDLPQIRSTLIPILHQKAKRGTPHQAKQAVH	CIHAIFTNKEVQLAQIFEP 720
Db	661	IFRNTGHKIE	TDLPQIRSTLIPILHQKAKRGTPHQAKQAVH	CIHAIFTNKEVQLAQIFEP 720
Qy	721	LSRSLNADVPEQLITPLVSLGHISMLAPDQFASPMKSVVANFIVKD	LLMNDRSTGEKNGK	780
Db	721	LSRSLNADVPEQLITPLVSLGHISMLAPDQFASPMKSVVANFIVKD	LLMNDRSTGEKNGK	780
Qy	781	LWSPDEEVSPEVLAKVQAIKLLVRWLLGMKNNQSKSANSTLRLLSAMLVSE	GDLTEQKRI	840
Db	781	LWSPDEEVSPEVLAKVQAIKLLVRWLLGMKNNQSKSANSTLRLLSAMLVSE	GDLTEQKRI	840
Qy	841	SKSDMSRLRLAAGSAIMKLAQEP	CYHEIITPEQFQLCALVINDECYQVRQIFAQKLHKAL	900
Db	841	SKSDMSRLRLAAGSAIMKLAQEP	CYHEIITPEQFQLCALVINDECYQVRQIFAQKLHKAL	900
Qy	901	VKLLLPLEYMAIFALCAKDPVKERRAHARQCLLNISIRREYIKQNP	MATEKLLSLLPEY	960
Db	901	VKLLLPLEYMAIFALCAKDPVKERRAHARQCLLNISIRREYIKQNP	MATEKLLSLLPEY	960
Qy	961	VVPYMIHLLAHDPDFTRSQD	VDQLRDIKECLWFMLEVLMTKNENNSHAFMCKMAENIKLT	1020
Db	961	VVPYMIHLLAHDPDFTRSQD	VDQLRDIKECLWFMLEVLMTKNENNSHAFMCKMAENIKLT	1020
Qy	1021	RDAQSPDESKTNEKLYTVC	DVALCVINSKSALCNADSPKDPVLP	MKFFTQPEKDFCNDKS 1080
Db	1021	RDAQSPDESKTNEKLYTVC	DVALCVINSKSALCNADSPKDPVLP	MKFFTQPEDKFCNDKS 1080
Qy	1081	YISEETRVLLLTGKPKPAGVLGAVNKPLSATGRKPYVRSTGTETGS	NINVNSELNPSTGN	1140
Db	1081	YISEETRVLLLTGKPKPAGVLGAVNKPLSATGRKPYVRSTGTETGS	NINVNSELNPSTGN	1140
Qy	1141	RSREQSSEAAETGVSENEENPVRIISVTPVKNIDPVKNKEINSDQATQGNIS	SDRGKKRT	1200
Db	1141	RSREQSSEAAETGVSENEENPVRIISVTPVKNIDPVKNKEINSDQATQGNIS	SDRGKKRT	1200
Qy	1201	VTAAGAENIQQKTDEK	VDESGPPAPSKPRRGRRPKSESQGNATKNDDL	NKPINKGRKRAA 1260
Db	1201	VTAAGAENIQQKTDEK	VDESGPPAPSKPRRGRRPKSESQGNATKNDDL	NKPINKGRKRAA 1260
Qy	1261	VGQESPGGLEAGNAKAPKLQDLAKKAAPAERQIDLQR		1297
Db	1261	VGQESPGGLEAGNAKAPKLQDLAKKAAPAERQIDLQR		1297

RESULT 4  
ABP64685  
ID ABP64685 standard; protein; 919 AA.  
XX

AC ABP64685;  
XX  
DT 25-FEB-2003 (first entry)  
XX  
DE Human protein SEQ ID 345.  
XX  
KW Human; expressed sequence tag; EST; haematopoietic disorder;  
KW central nervous system disease; viral infection;  
KW peripheral nervous system disease; non-healing wound; infectious disease;  
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;  
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;  
KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;  
KW cytostatic; haemostatic; virucide; antibacterial; fungicide;  
KW immunostimulant; cerebroprotective.  
XX  
OS Homo sapiens.  
XX  
PN WO200259260-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 16-NOV-2001; 2001WO-US042950.  
XX  
PR 17-NOV-2000; 2000US-00714936.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;  
PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
XX  
DR WPI; 2002-590824/63.  
DR N-PSDB; ABQ99271.  
XX  
PT New isolated polynucleotide, useful in research, diagnostic or  
PT therapeutic methods, e.g. preventing or treating disorders involving  
PT aberrant protein expression or biological activity.  
XX  
PS Claim 20; SEQ ID NO 345; 394pp; English.  
XX  
CC The present invention relates to novel human coding sequences (ABQ99268-  
CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in  
CC therapeutic, diagnostic and research methods. The polynucleotides may be  
CC used in the field of molecular biology as hybridisation probes, primers  
CC for PCR, for chromosome and gene mapping, for the recombinant production  
CC of protein, or in generation of anti-sense DNA or RNA. The  
CC polynucleotides are useful in diagnostics as expressed sequence tags  
CC (ESTs) for identifying expressed genes or for physical mapping of the  
CC human genome. The proteins may be used as molecular weight markers, or as  
CC nutritional sources or supplements. The proteins may be used to maintain  
CC and expand cell population in a totipotent or pluripotent state  
CC useful for re-engineering damaged or diseased tissues, transplantation,  
CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The  
CC polynucleotides and proteins are useful for preventing, treating or  
CC ameliorating disorders involving aberrant protein expression or  
CC biological activity, e.g. haematopoietic disorders, central/peripheral  
CC nervous system diseases, mechanical and traumatic disorders, non-healing  
CC wounds, immune deficiencies and disorders, infectious diseases caused by  
CC viral, bacterial or fungal infection, autoimmune disorders, allergic  
CC reactions and conditions, coagulation disorders, or cancer. The  
CC polynucleotide sequences of the invention were assembled from ESTs  
CC isolated mainly by sequencing by hybridisation, and in some cases,

CC sequences obtained from one or more public databases. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published pct sequences

XX

SO Sequence 919 AA;

Query Match 70.8%; Score 4711; DB 5; Length 919;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 919; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 379 MMGLAQLYKKYCLHGEAGKEAAEKVSWIKDKLLHIYYQNSIDDKLLVEKIFAQYLVPHNL 438

Db 1 MMGLAQLYKKYCLHGEAGKEAAEKVSWIKDKLLHIYYONSIDDKLLVEKIFAQYLVPHNL 60

Qy 439 ETEERMKCLYYLYASLD PNAVKALNEMWKCQNMLRSHVRELLDLHKQPTSEANC SAMFGK 498

Db 61 ETEERMKCLYYLYASLDPNAVKALNEMWKCONMLRSHVRELLDLHKOPTSEANC SAMFGK 120

QY            499. LMTIAKNLPDPGKAQDFVKKFNQVLGDDEKLRSQLELLISPTCSCKQADICVREIARKLA    558  
               |||||

Db 121 LMTIAKNLPDPGKAQDFVKKFNQVLGDDEKLRSQLELLISPTCSCKQADICVREIARKLA 180

Qy 559 NPKQPTNPFLEMVKFLLER IAPVHIDSEAI SALVKLMNKSIEGTADDEEEGVSPDTAIRS 618

Db 181 NPKQPTNPFLEMVKFLLERIAPVHIDSEAI SALVKLMNKSI EGTADDEEEGVSPDTAIRS 240

QY           619 GLELLKVLFSFTHPTSFHSAETYES LLQLRMEDDKVAEAAIQIFRNTGHKIETDLPQIRS 678  
| | | | |  
| | | | |

Db 241 GLELLKVLSFTHPTS FHS AETYESLLQCLRMEDDKVAEAAIQIFRNTGHK IETDLPQIRS 300

Qy 679 TLIPILHQAKRGTPHQAKQAVHC IHAIFTNKEVQLAQIFEPLSRSLNADVP EQLITPLV 738

Db 301 TLIPI LHQAKRGTPHQAKQAVHCIHAI FTNKEVQLAQIFEPLSRSLNADVPEQLITPLV 360

Qy           739 SLGHISMLAPDQFASPMKSVVANFIVKDLLMNDRSTGEKNGKLWSPDEEV SPEVLAKVQA       798  
| | | | |  
| | | | |

Db 361 SLGHI SMLAPDQFASPMKSVVANFIVKDLLMNDRSTGEKNGKLWSPDEEVSPPEVLAKVOA 420

Qy 799 IKLLVRWLLGMKNNQSKSANSTLRLLSAMLVSEGLTEQKRISKSDMSRLRLAAGSAIMK 858

Db 421 IKLLVRWLLGMKNNQSKSANSTLRLLSAMLVSEGLTEQKRISKSDMSRLRLAAGSAIMK 480

Qy           859 LAQEPCYHEIITPEQQFLCALVINDECYQVRQIFAQLHKALVKLLLPLEYMAIFALCAK   918  
| | | | | | | | | | | | | | | | | | | | | | | | | |

Db 481 LAQEP CYHEIITPEQFQLCALVINDECYQVRQIFAOKLHKALVKLLLPLEYMAIFALCAK 540

[illegible]

Db 541 DPVKERRAHARQCLLNISIRREYIKONPMATEKLLSLLPEYVVPYMIHLLAHDPPDFTRS 600

Qy 979 QDVDQLRDIKECLWFMLEVLMTKNENNSHAFMKKMAENIKLTRDAQSPDESKTNEKLYTV 1038

Db 601 QDVDQLRDIKECLWFMLEVLMTKNENNSHAFMKKMAENIKLTRDAOSPDESKTNEKLYTV 660

[illegible]

Db 661 CDVALCVINSKSA LCNADSPKDPVLP MKFFTQPEKDFCNDKSYISEETRVLLLTGKPKPA 720

Qy 1099 GVLGAVNKPLSATGRKPYVRSTGTETGSNI NVNSELNPSTGNRSREQSSEAAETGVSENE 1158

# SCORE Search Results Details for Application 10679580 and Search Result us-10-679-580a-2.rai.

[Score Home Page](#) [Retrieve Application List](#) [SCORE System Overview](#) [SCORE FAQ](#) [Comments / Suggestions](#)

This page gives you Search Results detail for the Application 10679580 and Search Result us-10-679-580a-2.rai.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2006, 21:11:59 ; Search time 55 Seconds  
(without alignments)  
2064.131 Million cell updates/sec

Title: US-10-679-580A-2  
Perfect score: 6650  
Sequence: 1 MIKRLKMVVKTFMDMDQDSE.....KLQDLAKKAAPAERQIDLQR 1297

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/5\_COMB.pep:\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/6\_COMB.pep:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/7\_COMB.pep:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/H\_COMB.pep:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/RE\_COMB.pep:\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	4668.5	70.2	1391	2	US-09-512-581D-2
2	4352.5	65.5	1346	2	US-09-949-016-9959

3	702	10.6	312	2	US-09-270-767-45520	Sequence 45520, A
4	377	5.7	92	2	US-09-513-999C-6892	Sequence 6892, Ap
5	374	5.6	309	2	US-09-270-767-45536	Sequence 45536, A
6	328.5	4.9	752	2	US-09-248-796A-15445	Sequence 15445, A
7	314	4.7	101	2	US-09-471-276-1408	Sequence 1408, Ap
8	284	4.3	235	2	US-09-270-767-43236	Sequence 43236, A
9	219	3.3	265	2	US-09-248-796A-15446	Sequence 15446, A
10	199	3.0	2663	2	US-09-538-092-1252	Sequence 1252, Ap
11	198	3.0	2662	2	US-09-595-684B-31	Sequence 31, Appl
12	195.5	2.9	3248	1	US-08-353-700-1	Sequence 1, Appli
13	195.5	2.9	3248	5	PCT-US95-16216-1	Sequence 1, Appli
14	191.5	2.9	2482	1	US-08-328-254-6	Sequence 6, Appli
15	190.5	2.9	3210	2	US-09-538-092-1154	Sequence 1154, Ap
16	189	2.8	1051	2	US-08-965-762-14	Sequence 14, Appl
17	189	2.8	1051	2	US-09-911-927-14	Sequence 14, Appl
18	189	2.8	1051	2	US-09-911-882-14	Sequence 14, Appl
19	189	2.8	1051	2	US-09-911-888-14	Sequence 14, Appl
20	181.5	2.7	169	2	US-09-270-767-33370	Sequence 33370, A
21	181.5	2.7	169	2	US-09-270-767-48587	Sequence 48587, A
22	181.5	2.7	2265	2	US-10-144-198-35	Sequence 35, Appl
23	180	2.7	2733	2	US-09-949-016-11433	Sequence 11433, A
24	177.5	2.7	3878	2	US-09-914-259-11	Sequence 11, Appl
25	177	2.7	2158	2	US-10-144-198-34	Sequence 34, Appl
26	177	2.7	2697	2	US-10-144-198-12	Sequence 12, Appl
27	176.5	2.7	1979	2	US-09-949-016-6468	Sequence 6468, Ap
28	176.5	2.7	2047	2	US-09-949-016-7404	Sequence 7404, Ap
29	176	2.6	1055	2	US-09-949-016-9776	Sequence 9776, Ap
30	175	2.6	3696	2	US-09-134-001C-5080	Sequence 5080, Ap
31	174	2.6	3259	2	US-09-949-016-6507	Sequence 6507, Ap
32	172	2.6	994	2	US-09-949-016-6779	Sequence 6779, Ap
33	172	2.6	1173	2	US-09-248-796A-19313	Sequence 19313, A
34	168	2.5	2704	2	US-09-538-092-1260	Sequence 1260, Ap
35	167.5	2.5	2954	2	US-09-150-867-1	Sequence 1, Appli
36	167.5	2.5	2954	3	US-09-724-584-1	Sequence 1, Appli
37	164.5	2.5	2101	1	US-08-466-390-4	Sequence 4, Appli
38	164.5	2.5	2101	1	US-08-470-950-4	Sequence 4, Appli
39	164.5	2.5	2101	1	US-08-467-781-4	Sequence 4, Appli
40	164.5	2.5	2101	1	US-08-195-487-4	Sequence 4, Appli
41	164.5	2.5	2101	1	US-08-483-924-4	Sequence 4, Appli
42	164.5	2.5	2101	2	US-09-452-294-1	Sequence 1, Appli
43	164.5	2.5	2101	2	US-09-296-662-32	Sequence 32, Appl
44	164.5	2.5	2101	5	PCT-US93-06160-4	Sequence 4, Appli
45	164.5	2.5	2115	2	US-09-296-662-33	Sequence 33, Appl

## ALIGNMENTS

## RESULT 1

US-09-512-581D-2

; Sequence 2, Application US/09512581D

; Patent No. 6994992

; GENERAL INFORMATION:

; APPLICANT: Soto, Ana M.

; APPLICANT: Sonnenschein, Carlos

; APPLICANT: Geck, Peter

; APPLICANT: Szelei, Jozsef

; TITLE OF INVENTION: NOVEL ANDROGEN-INDUCED SUPPRESSOR OF CELL PROLIFERATION AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 34724-026

; CURRENT APPLICATION NUMBER: US/09/512,581D

; CURRENT FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: US 60/121,461  
 ; PRIOR FILING DATE: 1999-02-24  
 ; NUMBER OF SEQ ID NOS: 72  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 2  
 ; LENGTH: 1391  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-512-581D-2

Query Match 70.2%; Score 4668.5; DB 2; Length 1391;  
 Best Local Similarity 69.0%; Pred. No. 0;  
 Matches 908; Conservative 160; Mismatches 199; Indels 49; Gaps 11;

Qy	1	MIKRLKMVVKT	FMDMDQDSEDEKQ	QYLPLALHLASE	FFLRNP	NKDVRL	LVACCLADIFRI	60
		::						
Db	31	MVRLKMVVKT	FMDMDQDSEEEKEL	YNLALHLASD	FFLKHPGK	DVRL	LVACCLADIFRI	90
Qy	61	YAPEAPYTSHDKL	KDIFLFITRQL	KGLEDTKSPQFNRY	FYLLEN	LAWVKS	SYNICFELEDC	120
Db	91	YAPEAPYTS	PDKLKDI	FMFITRQL	KGLEDTKSPQFNRY	FYLLEN	IAWVKS	150
Qy	121	NEIFIQLFRTL	FSVINNSHNK	KVQMHMLDL	MSSI	IMEGDGVT	QELLGSILINLI	180
Db	151	NEIFTQLYRTL	FSVINNGHNQ	KVHMHMVDL	MSSI	ICEGDTV	SQELLDTVL	210
Qy	181	NKQSFDAKVL	LKRTVQTI	EACIANFFNQ	VLVGRSSV	SDLSEHV	FDLIQELFAIDPHLL	240
Db	211	NKQAYDLAK	ALLKRTAQA	IPEYITTF	FNQVLM	LKGT	SI	270
Qy	241	LSVMPQLEF	KLKSN	DGEERL	AVRLLAK	LFGSKD	SLATQNRPLWQ	300
Db	271	LSVLPQLEF	KLKSN	DNEERL	QVVKLLAK	MFGAKD	SELASQNKPLWQ	330
Qy	301	LESVKFASH	CLMNHPDL	AKDLTEY	LKVRSHD	PEEAIR	RHDVIVTI	360
Db	331	LECVKFASH	CLMNHPDL	AKDLTEY	LKVRSHD	PEEAIR	RHDVIVSI	390
Qy	361	GFVRERTL	DKRWRVR	KEAMMGLA	QLYKKYCL	HGEAGKEA	AEKVSWIKDKLLH	420
Db	391	NFVRERTL	DKRWRVR	KEAMMGLA	QIYKKYAL	QSAAGK	DAKQIAWIKDKLLH	450
Qy	421	DKLLVEKIF	AQYLVPHN	LETEERM	KCLYYLYA	SLDPNAV	KALNEMWKCQNMLR	480
Db	451	DRLLVERIF	AQYMPVPHN	LETTERM	KCLYYLYA	TDLNAV	KALNEMWKCQNLLRHQ	510
Qy	481	DLHKQPT	SEANCSAM	FGKLMT	IAKNLP	DPGKAQ	DFVKKFNQVLG	540
Db	511	DLIKQPKT	DASVKAI	FSKVMVIT	RNLPDP	GKAQ	DFMCKFTQVLE	570
Qy	541	CSCQADIC	VREIARK	LANPKQPT	NPFLEM	VKFLLE	RIAPVHIDSEAI	600
Db	571	CSCQAE	GCVREIT	TKKLG	NPKQPT	NPFLEM	IKFLLE	630
Qy	601	GTADDEEE	GVSPDTA	IRSGLELL	KVLSF	THPTSF	HSAETYESLLQ	660
Db	631	GTADDE	DEGVPTD	QAIRAG	LELLK	VLSF	THPISF	690
Qy	661	IFRNTGH	KIETDLP	QIRSTLI	PILHQ	AKRGT	PHQAKQAVH	720

Db	691	IFKNTGSKIEEDFPHIRSALLPVLHHKSKKGPPRQAKYAIHCIHAIFSSKETQFAQIFEP	750
Qy	721	LSRSLNADVPEQLITPLVSLGHISMLAPDQFASPMKSVVANFIVKDLLMNRSTGEKNGK	780
		:  :        : : : : : :                :	
Db	751	LHKSLDPSNLEHLITPLVTIGHIALAPDQFAAPWKSWSVATFIVKDLLMNDRLPGKKTTK	810
Qy	781	LWSPDEEVSPPEVLAKVQAIKLLVRWLLGMKNQSKSANSTLRLLSAMLVSEGDLTEQKRI	840
		:   :   : : : : :              :     :      :	
Db	811	LWVPDEEVSPETMVKIQAIKMMVRWLLGMKNHNSKSGTSTLRLLTTILHSDGDLTEQGKI	870
Qy	841	SKSDMSRLRLAAGSAIMKLAQEPYHEIITPEQFQCALVINDECYQVRQIFAQKLHKAL	900
		:   :   : : : : :                :	
Db	871	SKPDMSRLRLAAGSAIVKLAQEPYHEIITLEQYQLCALAINDECYQVRQVFAQKLHKGL	930
Qy	901	VKLLLPLEYMAIFALCAKDPVKERRAHARQCLLKNI SIRREYIKQNPMA TEKLLSLLPEY	960
		:                   :   : : : : :  :         :	
Db	931	SRLRLPLEYMAICALCAKDPVKERRAHARQCLVKNI NVRREYLKQHA AVSEKLLSLLPEY	990
Qy	961	VVPYMIHLLAHDPDFTRSQDQDLRDIKECLWFMLEVLMTKNENNSHAFMKKMAENIKLT	1020
		: :   : : : : : :              :	
Db	991	VVPYTIHLLAHDPDYVKVDIEQLKDVKECLWFVLEILMAKNENNSHAFIRKMVENIKQT	1050
Qy	1021	RDAQSPDESKTNEKLYTVCDVALCVINSKALCNADSPKDPVLPMKFFTQPEKDFCNDKS	1080
		:         :          :  :       :   :       :   :       :	
Db	1051	KDAQGPDDAKMNEKLYTVCDVAMNIIMSKSTTYSLESPKDPVLPARFFTQPDKNFSNTKN	1110
Qy	1081	YISEETRVLLLTGKPKPAGVLGAVNKPLSATGRKPYVRSTGTETGSNINVNSELNPSTGN	1140
		:   :             :   : :   :         :       :	
Db	1111	YLPPEMKSFFTPGKPKTTNVLGAVNKPLSSAGKQSQTSSRMETVS--NASSSSNPSSPG	1168
Qy	1141	RSREQSSEAAETGVSENEENPVRIISVTPVKKNID-----PVKNK	1179
		: : : :       : :	
Db	1169	RIKGR-LDSSEMDHSENE--YTMSSPLPGKSKDKRDDSDLVRSELEKPRGRKKTPTVTEQ	1225
Qy	1180	E--INSDQAT----QGNISSDRGKKRTVTAAGAENIQQKTDEK-----VDESGPP	1223
		:             :       : :     :	
Db	1226	EEKLGMDDLTKLVQEQKPKGSQRSRKRGTASESDE-QQWPEEKRLKEDILENEDEQNSP	1284
Qy	1224	APSKPRRGRRPKSESQGNATKNDDL NKPINKGRKR---AAVGQESPGGLEAGNAK	1275
		:             :       :     :   :   :	
Db	1285	-PKKGKRGRPPKPLG-GGT PKEEPTMKTSKKGSKKSGPPAPEEEEEERQSGNTE	1338

RESULT 2  
US-09-949-016-9959  
; Sequence 9959, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012



[Score Home Page](#) [Retrieve Application List](#) [SCORE System Overview](#) [SCORE FAQ](#) [Comments / Suggestions](#)

This page gives you Search Results detail for the Application 10679580 and Search Result us-10-679-580a-2.rpr.  
[start](#)

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2006, 21:06:29 ; Search time 53 Seconds  
(without alignments)  
2354.588 Million cell updates/sec

Title: US-10-679-580A-2  
Perfect score: 6650  
Sequence: 1 MIKRLKMVVKTFMDMDQDSE.....KLQDLAKKAAPAERQIDLQR 1297

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	IDB	ID	Description
1	4350	65.4	851	2	T00374	hypothetical prote
2	661.5	9.9	1579	2	T23142	hypothetical prote
3	652	9.8	1506	2	S52957	bimD protein - Eme
4	621.5	9.3	1303	2	E96805	hypothetical prote
5	538	8.1	1277	2	S54451	hypothetical prote
6	257	3.9	390	2	T43647	hypothetical prote
7	256.5	3.9	780	2	F96840	hypothetical prote
8	250	3.8	990	2	H86293	protein T24D18.4 [

9	220	3.3	852	2	T06310	hypothetical prote
10	214.5	3.2	1875	2	S38173	myosin-like protei
11	204.5	3.1	1790	2	S67593	transport protein
12	203	3.1	1819	2	A71928	cag island protein
13	200.5	3.0	2139	2	T18296	myosin heavy chain
14	199	3.0	2663	1	S28261	centromere protein
15	198.5	3.0	1927	2	G64585	cag pathogenicity
16	196.5	3.0	1160	2	T13713	beta3 protein - fr
17	192	2.9	1679	2	S48385	hypothetical prote
18	191.5	2.9	1302	1	JC6009	surface-located me
19	190.5	2.9	1916	2	S46157	RIF1 protein - yea
20	190	2.9	2253	2	T30336	nuclear/mitotic ap
21	189	2.8	1051	2	S59791	probable membrane
22	188	2.8	1630	2	S64403	ESP1 protein - yea
23	185.5	2.8	4385	2	T29042	hypothetical prote
24	185	2.8	2297	2	AB2494	hypothetical prote
25	183.5	2.8	1721	2	T21214	hypothetical prote
26	182.5	2.7	2819	2	A90551	conserved hypothet
27	180.5	2.7	932	2	S65214	probable alpha/gam
28	180	2.7	3225	2	I52300	giantin - human
29	179.5	2.7	2346	2	T13829	Tpr homolog - frui
30	179	2.7	2116	2	A26655	myosin heavy chain
31	178	2.7	1034	2	T17458	chromosome condens
32	175.5	2.6	1642	2	T08880	NMDA receptor-bind
33	175.5	2.6	2493	2	S45734	probable membrane
34	174	2.6	2692	2	T23768	hypothetical prote
35	174	2.6	3259	1	A56539	giantin - human
36	173.5	2.6	1931	2	A59234	slow myosin heavy
37	173	2.6	798	2	I50479	neurofilament medi
38	173	2.6	1226	2	I51617	kinesin-like prote
39	172.5	2.6	1837	2	T41023	probable nuclear p
40	172	2.6	845	2	A45669	neurofilament trip
41	172	2.6	1727	2	T50073	myosin-like coiled
42	172	2.6	2481	2	D90011	FmtB protein [impo
43	170.5	2.6	1447	2	F82909	hypothetical prote
44	170.5	2.6	1864	2	F86378	protein F21J9.12 [
45	170.5	2.6	2415	1	A33733	spectrin alpha cha

## ALIGNMENTS

## RESULT 1

T00374

hypothetical protein KIAA0648 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 09-Jul-2004

C;Accession: T00374

R;Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O  
DNA Res. 5, 169-176, 1998

A;Title: Prediction of the coding sequences of unidentified human genes. X. The complete sequences

A;Reference number: Z14142; MUID:98403880; PMID:9734811

A;Accession: T00374

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-851

A;Cross-references: UNIPROT:Q9Y4D4; UNIPARC:UPI000006D34B; EMBL:AB014548; NID:g3327109; PIDN:BAA31

A;Experimental source: brain

C;Genetics:

A;Note: KIAA0648

http://es/ScoreAccessWeb/GetItem.action?AppId=10679580&seqId=567837&ItemName=us-10-679-580a-... 8/6/2006

```

RESULT 2
T23142
hypothetical protein H38K22.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T23142; T24139
R;Barlow, K.
submitted to the EMBL Data Library, June 1998
A;Reference number: Z19695
A;Accession: T23142
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1579
A;Cross-references: UNIPROT:Q9XTF6; UNIPARC:UPI0000078215; EMBL:AL024499; PIDN:CAA19710.1; GSPDB:GI
A;Experimental source: clone H38K22
R;Ainscough, R.
submitted to the EMBL Data Library, August 1995
A;Reference number: Z19843
A;Accession: T24139
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1579
A;Cross-references: UNIPARC:UPI0000078215; EMBL:Z50874; PIDN:CAA90774.1; GSPDB:GN00021; CESP:H38K2
A;Experimental source: clone R10E4
C;Genetics:
A;Gene: CESP:H38K22.1
A;Map position: 3
A;Introns: 28/3; 99/3; 130/3; 239/3; 323/3; 363/2; 396/1; 444/3; 559/2; 651/3; 689/3; 793/1; 831/3

```

Qy	2	IKRLKMVVKTFMDMDQDSEDE---KQYYLPLALHLASEFFLRNPKNKDVRLLVACCLADIF	58
Db	24	IERLRKLFNCLHDCKTNNSEEVSSPNRFARLFQHLSQECFLDNSNTDFRILLSLCLANIL	83
Qy	59	RIYAPEAPYTSHTDKLDIFLFITRQLKGLEDT--KSPQFNRYFYLLLENLAWVKSYNICFE	116
Db	84	RIFQPELPTPSVMDLKEVYIYLFRTMRGLGDVTQDSPKFKNYFSLVETME--KIIPPIIE	141
Qy	117	LEDCNE-----IFIQLFRTLFSV-----INNShN-----	140
Db	142	MKDHDDKEATPVFRALIKDILAIPCGKGWNQNLKKEARLLKIQENDDDSMNDDEEDENAA	201
Qy	141	KKVQMHMLDLMSIIMEGDGVTQELLGSILINLI PAHKNLNKQSFDLAKVLLKRTVQTIE	200
Db	202	EKIRKSLIQIATTVITNLD FVQNECLDVLFYHII---NPQRSNFABEARALAE---DIIR	254
Qy	201	ACIANFFNQVLVLGRSSVSDLSEH-----VFDLIQELFAIDPHLLLSV	243
Db	255	SCSDNESDTLANSIRSTMTAAAKEGKLPEEFELTGSSNRSKFFEVRLYLHYVSFDLVSGA	314
Qy	244	MPQLEFKLKSNDGEERLAVVRLLA KLFGSKDSDLA-TQNRPLWQCFLGRFNDIHVPVRLE	302
Db	315	IOELKFWLOSENEYRKEAVTVVGMLTRDKHCOFGMDSNDPTWSAFLNASIDODDSVRHE	374

# SCORE Search Results Details for Application 10679580 and Search Result us-10-679-580a-2.rup.

[Score Home Page](#) [Retrieve Application List](#) [SCORE System Overview](#) [SCORE FAQ](#) [Comments / Suggestions](#)

This page gives you Search Results detail for the Application 10679580 and Search Result us-10-679-580a-2.rup.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2006, 21:02:48 ; Search time 327 Seconds  
(without alignments)  
3668.946 Million cell updates/sec

Title: US-10-679-580A-2  
Perfect score: 6650  
Sequence: 1 MIKRLKVVVKTFFMDMDQDSE.....KLQDLAKKAAPAERQIDLQR 1297

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_7.2:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6650	100.0	1297	2	Q8NG14_HUMAN	Q8ng14 homo sapien
2	6643	99.9	1297	2	Q68DF7_HUMAN	Q68df7 homo sapien
3	6232.5	93.7	1356	2	Q5F3V3_CHICK	Q5f3v3 gallus gall
4	5851	88.0	1323	2	Q4KLU7_XENLA	Q4klu7 xenopus lae
5	5849	88.0	1323	2	Q4QXM3_XENLA	Q4qxm3 xenopus lae
6	5462	82.1	1122	2	Q6A026_MOUSE	Q6a026 mus musculu
7	4694.5	70.6	1446	2	Q4VA53_MOUSE	Q4va53 mus musculu

8	4682.5	70.4	1447	2	Q9NTI5_HUMAN	Q9nti5	homo sapien
9	4682.5	70.4	1483	2	Q9Y2I5_HUMAN	Q9y2i5	homo sapien
10	4678	70.3	1262	2	Q5F3U9_CHICK	Q5f3u9	gallus gall
11	4675.5	70.3	1446	2	Q7TSS4_MOUSE	Q7tss4	mus musculu
12	4652.5	70.0	1391	2	Q9Y451_HUMAN	Q9y451	homo sapien
13	4638.5	69.8	1448	2	Q498H0_XENLA	Q498h0	xenopus lae
14	4593	69.1	1199	2	Q5U241_XENLA	Q5u241	xenopus lae
15	4350	65.4	851	2	Q9Y4D4_HUMAN	Q9y4d4	homo sapien
16	4126.5	62.1	1438	2	Q4SG13_TETNG	Q4sg13	tetraodon n
17	3712.5	55.8	1191	2	Q80TM8_MOUSE	Q80tm8	mus musculu
18	3686	55.4	912	2	Q5R3S3_HUMAN	Q5r3s3	homo sapien
19	2946	44.3	584	2	Q8VDS0_MOUSE	Q8vds0	mus musculu
20	2842	42.7	600	2	Q8N7J4_HUMAN	Q8n7j4	homo sapien
21	2838	42.7	600	2	Q2TTR5_HUMAN	Q2ttr5	homo sapien
22	2612.5	39.3	965	2	Q4QXM2_XENLA	Q4qxm2	xenopus lae
23	2456.5	36.9	701	2	Q8BJ18_MOUSE	Q8bj18	mus musculu
24	2118.5	31.9	824	2	Q8BLH6_MOUSE	Q8blh6	mus musculu
25	2034	30.6	529	2	Q8IXT6_HUMAN	Q8ixt6	homo sapien
26	2033	30.6	454	2	Q4T0S9_TETNG	Q4t0s9	tetraodon n
27	2030.5	30.5	1130	2	Q7PRK5_ANOGA	Q7prk5	anopheles g
28	1911	28.7	1218	2	Q9V646_DROME	Q9v646	drosophila
29	1744.5	26.2	432	2	Q7SXH3_BRARE	Q7sxh3	brachydanio
30	1715	25.8	333	2	Q96DB6_HUMAN	Q96db6	homo sapien
31	1047	15.7	242	2	Q6TRW4_RAT	Q6trw4	rattus norv
32	999	15.0	198	2	Q9D337_MOUSE	Q9d337	mus musculu
33	997.5	15.0	533	2	Q5W0K8_HUMAN	Q5w0k8	homo sapien
34	925.5	13.9	202	2	Q4SZB7_TETNG	Q4szb7	tetraodon n
35	856.5	12.9	1561	2	Q5VMT5_ORYSA	Q5vmt5	oryza sativ
36	785.5	11.8	1250	2	Q4P823_USTMA	Q4p823	ustilago ma
37	779	11.7	1205	1	PDS5_SCHPO	Q9hff5	schizosacch
38	769.5	11.6	1638	2	Q9FIL0_ARATH	Q9fil0	arabidopsis
39	725.5	10.9	1279	2	Q55KR8_CRYNE	Q55kr8	cryptococcu
40	725.5	10.9	1279	2	Q5KAQ6_CRYNE	Q5kaq6	cryptococcu
41	701.5	10.5	1481	2	Q2UB28_ASPOR	Q2ub28	aspergillus
42	696	10.5	1690	2	Q60YT4_CAEBR	Q60yt4	caenorhabdi
43	690.5	10.4	1523	2	Q4I755_GIBZE	Q4i755	gibberella
44	685.5	10.3	1492	2	Q4WZI1_ASPFU	Q4wzi1	aspergillus
45	673.5	10.1	1596	2	Q9UVY6_SORMA	Q9uvy6	sordaria ma

## ALIGNMENTS

## RESULT 1

Q8NG14\_HUMAN

ID Q8NG14\_HUMAN PRELIMINARY; PRT; 1297 AA.

AC Q8NG14;

DT 01-OCT-2002, integrated into UniProtKB/TrEMBL.

DT 01-OCT-2002, sequence version 1.

DT 07-FEB-2006, entry version 13.

DE SCC-112.

GN Name=SCC-112;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX PubMed=15019998; DOI=10.1016/j.gene.2003.12.013;

RA Kumar D., Sakabe I., Patel S., Zhang Y., Ahmad I., Gehan E.A.,

RA Whiteside T.L., Kasid U.;  
 RT "SCC-112, a novel cell cycle-regulated molecule, exhibits reduced  
 RT expression in human renal carcinomas.";   
 RL Gene 328:187-196 (2004).  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AF294791; AAM82347.1; -; mRNA.  
 DR Ensembl; ENSG00000121892; Homo sapiens.  
 DR GO; GO:0005488; F:binding; IEA.  
 DR InterPro; IPR011989; ARM-like.  
 DR InterPro; IPR000357; HEAT.  
 DR Pfam; PF02985; HEAT; 1.  
 SQ SEQUENCE 1297 AA; 146610 MW; A42D13E68CA5808E CRC64;

Query Match 100.0%; Score 6650; DB 2; Length 1297;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MIKRLKMVVKTFMDMDQDSEDEKQYLPALHLASEFFLRNPNDVRLLVACCLADIFRI	60
Db	1	MIKRLKMVVKTFMDMDQDSEDEKQYLPALHLASEFFLRNPNDVRLLVACCLADIFRI	60
Qy	61	YAPEAPYTS HDKLDIFLFITRQLKGLEDTKSPQFNRYFYLLLENLAWVKSYNICFELEDC	120
Db	61	YAPEAPYTS HDKLDIFLFITRQLKGLEDTKSPQFNRYFYLLLENLAWVKSYNICFELEDC	120
Qy	121	NEIFIQLFRTLFSVINNSHNKKVQMHLMDLMSIIMEGDGVTQELLGSILINLI PAHKNL	180
Db	121	NEIFIQLFRTLFSVINNSHNKKVQMHLMDLMSIIMEGDGVTQELLGSILINLI PAHKNL	180
Qy	181	NKQSFDLAKVLLKRTVQTIEACIANFFNQVLVLGRSSVSDLSEHVFDLIQELFAIDPHLL	240
Db	181	NKQSFDLAKVLLKRTVQTIEACIANFFNQVLVLGRSSVSDLSEHVFDLIQELFAIDPHLL	240
Qy	241	LSVMPQLEFKLKSNDGEERLAVRLLAKLFGSKSDLATQNRPLWQCFLGRFNDIHVPVR	300
Db	241	LSVMPQLEFKLKSNDGEERLAVRLLAKLFGSKSDLATQNRPLWQCFLGRFNDIHVPVR	300
Qy	301	LESVKFASHCLMNHPDLAKDLTEYLKVRSHDPEEAI RHDVIVTIITAAKRDALVNDQLL	360
Db	301	LESVKFASHCLMNHPDLAKDLTEYLKVRSHDPEEAI RHDVIVTIITAAKRDALVNDQLL	360
Qy	361	GFVRERTL DKRWRVRKEAMMGLAQLYKKYCLHGEAGKEAAEKVSWIKDKLLHIYYQNSID	420
Db	361	GFVRERTL DKRWRVRKEAMMGLAQLYKKYCLHGEAGKEAAEKVSWIKDKLLHIYYQNSID	420
Qy	421	DKLLVEKIFAQYLVPHNLETEERMKCLYYLYASLDPNVAVKALNEMWKCQNMLRSHVRELL	480
Db	421	DKLLVEKIFAQYLVPHNLETEERMKCLYYLYASLDPNVAVKALNEMWKCQNMLRSHVRELL	480
Qy	481	DLHKQPTSEANCSAMFGKLM TI AKNLPDPGKAQDFVKKFNQVLGDDEKLRSQLELLISPT	540
Db	481	DLHKQPTSEANCSAMFGKLM TI AKNLPDPGKAQDFVKKFNQVLGDDEKLRSQLELLISPT	540
Qy	541	CSCQADICVREIARKLANPKQPTNP FLEMVKFLLERIAPVHIDSEAI SALVKLMNKSIE	600
Db	541	CSCQADICVREIARKLANPKQPTNP FLEMVKFLLERIAPVHIDSEAI SALVKLMNKSIE	600
Qy	601	GTADDEEEGVSPDTAIRSGLELLKVL SFTHPTSFHSAETYESLLQCLRMEDDKVAEAAIQ	660

Db	601	GTADDEEEGVSPDTAIRSGLELLKVLSFTHPTSFHSAETYESLLQCLRMEDDKVAEAAIQ	660
Qy	661	IFRNTGHKIEIDLPIRSTLIPILHQKAKRGTPHQAKQAVHCIHAIFTNKEVQLAQIFEP	720
Db	661	IFRNTGHKIEIDLPIRSTLIPILHQKAKRGTPHQAKQAVHCIHAIFTNKEVQLAQIFEP	720
Qy	721	LSRSLNADVPEQLITPLVSLGHISMLAPDQFASPMKSVVANFIVKDLLMNDRSTGEKNGK	780
Db	721	LSRSLNADVPEQLITPLVSLGHISMLAPDQFASPMKSVVANFIVKDLLMNDRSTGEKNGK	780
Qy	781	LWSPDEEVSPEVLAKVQAIKLLVRWLLGMKNNQSKSANSTLRLLSAMLVSEGDLEQKRI	840
Db	781	LWSPDEEVSPEVLAKVQAIKLLVRWLLGMKNNQSKSANSTLRLLSAMLVSEGDLEQKRI	840
Qy	841	SKSDMSRLRLAAGSAIMKLAQEPYHEIITPEQFQLCALVINDECYQVRQIFAQKLHKAL	900
Db	841	SKSDMSRLRLAAGSAIMKLAQEPYHEIITPEQFQLCALVINDECYQVRQIFAQKLHKAL	900
Qy	901	VKLLLPLEYMAIFALCAKDPVKERRAHARQCLLKNI SIRREYIKQNPMA TEKLLSLLPEY	960
Db	901	VKLLLPLEYMAIFALCAKDPVKERRAHARQCLLKNI SIRREYIKQNPMA TEKLLSLLPEY	960
Qy	961	VVPYMIHLLAHDPDFTRSQDQDQLRDIKECLWFMLEVLMTKNENNSHAFMKKMAENIKLT	1020
Db	961	VVPYMIHLLAHDPDFTRSQDQDQLRDIKECLWFMLEVLMTKNENNSHAFMKKMAENIKLT	1020
Qy	1021	RDAQSPDESKTNEKLYTVCDVALCVINSKSALCNADSPKDPVLPMKFFTQPEKDFCNDKS	1080
Db	1021	RDAQSPDESKTNEKLYTVCDVALCVINSKSALCNADSPKDPVLPMKFFTQPEKDFCNDKS	1080
Qy	1081	YISEETRVLLL TGKPKPAGVLGAVNKPLSATGRKPYVRSTGTETGSNINVNSELNPSTGN	1140
Db	1081	YISEETRVLLL TGKPKPAGVLGAVNKPLSATGRKPYVRSTGTETGSNINVNSELNPSTGN	1140
Qy	1141	RSREQSSEAAETGVSENEENPVRIISVTPVKNI DPVKNKEINSDQATQGNISSDRGKKRT	1200
Db	1141	RSREQSSEAAETGVSENEENPVRIISVTPVKNI DPVKNKEINSDQATQGNISSDRGKKRT	1200
Qy	1201	VTAAGAENIQKKTDEKVDSESGPPAPSKPRGRRPKSESQGNATKNDDL NKPINKGRKRAA	1260
Db	1201	VTAAGAENIQKKTDEKVDSESGPPAPSKPRGRRPKSESQGNATKNDDL NKPINKGRKRAA	1260
Qy	1261	VGQESPGGLEAGNAKAPKLQDLAKKAAPAERQIDLQR	1297
Db	1261	VGQESPGGLEAGNAKAPKLQDLAKKAAPAERQIDLQR	1297

## RESULT 2

## Q68DF7\_HUMAN

ID Q68DF7\_HUMAN PRELIMINARY; PRT; 1297 AA.

AC Q68DF7;

DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.

DT 11-OCT-2004, sequence version 1.

DT 07-FEB-2006, entry version 12.

DE Hypothetical protein DKFZp686B19246.

GN Name=DKFZp686B19246;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.



OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Salivary gland;  
 RG The German cDNA Consortium;  
 RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,  
 RA Fobo G., Han M., Wiemann S.;  
 RL Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; CR749425; CAH18263.1; -; mRNA.  
 DR Ensembl; ENSG00000121892; Homo sapiens.  
 DR GO; GO:0005488; F:binding; IEA.  
 DR InterPro; IPR011989; ARM-like.  
 DR InterPro; IPR000357; HEAT.  
 DR Pfam; PF02985; HEAT; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 1297 AA; 146668 MW; 3F6FC3D79DCE238D CRC64;

Query Match 99.9%; Score 6643; DB 2; Length 1297;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1296; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MIKRLKMOVVKT	FMDMDQDSEDEKQY	LPLALHLASEFFLRNPNKDVRL	LVACCLADIFRI	60
Db	1	MIKRLKMOVVKT	FMDMDQDSEDEKQY	LPLALHLASEFFLRNPNKDVRL	LVACCLADIFRI	60
Qy	61	YAPEAPYTSHDKL	KDIFLFITRQLK	GLEDTKSPQFNRYFY	LLENLAWVKS	120
Db	61	YAPEAPYTSHDKL	KDIFLFITRQLK	GLEDTKSPQFNRYFY	LLENLAWVKS	120
Qy	121	NEIFIQLFRTLFS	VINNSHNKKVQ	MHMLDLMS	SIIMEGDGVTQELL	180
Db	121	NEIFIQLFRTLFS	VINNSHNKKVQ	MHMLDLMS	SIIMEGDGVTQEL	180
Qy	181	NKQSFDLAKVLL	KKRTVQTI	EACIANFFNQVL	VLRSSVSDLSEHV	240
Db	181	NKQSFDLAKVLL	KKRTVQTI	EACIANFFNQVL	VLRSSVSDLSEHV	240
Qy	241	LSVMPQLEFKL	KSNDGEERL	AVRLLAKLFG	SKSDLATQNRPL	300
Db	241	LSVMPQLEFKL	KSNDGEERL	AVRLLAKLFG	SKSDLATQNRPL	300
Qy	301	LESVKFASHCL	MNHPDLAKDL	TEYLKVRSHD	PEEAIRHDVIVTI	360
Db	301	LESVKFASHCL	MNHPDLAKDL	TEYLKVRSHD	PEEAIRHDVIVTI	360
Qy	361	GFVRERTLDKR	WRVRKEAM	MGLAQLYKKY	CLHGEAGKEAAEK	420
Db	361	GFVRERTLDKR	WRVRKEAM	MGLAQLYKKY	CLHGEAGKEAAEK	420
Qy	421	DKLLVEKIFAQ	YLVPHNLE	TEERMKCLY	YLYASLDPNAV	480
Db	421	DKLLVEKIFAQ	YLVPHNLE	TEERMKCLY	YLYASLDPNAV	480
Qy	481	DLHKQPTSEAN	CSAMFGKL	MTIAKNLPD	PGKAQDFVKKFN	540
Db	481	DLHKQPTSEAN	CSAMFGKL	MTIAKNLPD	PGKAQDFVKKFN	540

```

RESULT 3
Q5F3V3_CHICK
ID   Q5F3V3_CHICK    PRELIMINARY;      PRT;   1356 AA.
AC   Q5F3V3;
DT   15-MAR-2005, integrated into UniProtKB/TrEMBL.
DT   15-MAR-2005, sequence version 1.
DT   07-FEB-2006, entry version 4.
DE   Hypothetical protein.

```

GN ORFNames=RCJMB04\_6f4;  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=CB; TISSUE=Bursa;  
 RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,  
 RA Fiedler P., Kutter S., Blagodatski A., Kostovska D., Koter M.,  
 RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;  
 RT "Full-length cDNAs from chicken bursal lymphocytes to facilitate  
 RT gene function analysis."  
 RL Genome Biol. 6:R6-R6(2005).  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AJ851547; CAH65181.1; -; mRNA.  
 DR GO; GO:0005488; F:binding; IEA.  
 DR InterPro; IPR011989; ARM-like.  
 DR InterPro; IPR000357; HEAT.  
 DR Pfam; PF02985; HEAT; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 1356 AA; 153085 MW; 4C9D9F2FB6DA9C42 CRC64;

Query Match 93.7%; Score 6232.5; DB 2; Length 1356;  
 Best Local Similarity 93.5%; Pred. No. 0;  
 Matches 1213; Conservative 36; Mismatches 46; Indels 3; Gaps 3;

Qy	1	MIKRLKMVVKTFMDMDQDSEDEKQQYLPLALHLASEFFLRNPNKDVRLLVACCLADIFRI	60
		: :	
Db	61	VVKRLKMVVKTFMDMDQDSEDEKQQYLPLALHLASEFFLRNPNKDVRLLVACCLADIFRI	120
Qy	61	YAPEAPYTS HDKLKDIFLFITRQLKGLEDTKSPQFNRYFYLLLENLAWVKSYNICFELEDC	120
Db	121	YAPEAPYTS HDKLKDIFLFITRQLKGLEDTKSPQFNRYFYLLLENLAWVKSYNICFELEDC	180
Qy	121	NEIFIQLFRTLFSVINNSHNKKVQMHMLDLMSIIMEGDGVTQELLGSILINLI PAHKNL	180
		:	
Db	181	NEIFIQLFRTLFSVINNSHNQKVQMHMLDLMSIIMEGDGVTQELLDSILINLI PAHKNL	240
Qy	181	NKQSFDLAKVLLKRTVQTIEACIANFFNQVLVLGRSSVSDLSEHVFDLIQELFAIDPHLL	240
		:     :	
Db	241	NKQAFDLAKVLLKRTVQTIEPCIANFFNQVLVLGKSSVSDLSEHVFDLILELFAIDPHLL	300
Qy	241	LSVMPQLEFKLKSNDGEERLAVRLLAKLFGSKSDLATQNRPLWQCFLGRFNDIHVPVR	300
Db	301	LSVMPQLEFKLKSNDGEERLAVRLLAKLFGSKSDLATQNRPLWQCFLGRFNDIHVPVR	360
Qy	301	LESVKFASHCLMNHPDLAKDLTEYLKVRSHDPEEAI RHDVIVTIITAAKRDALVNDQLL	360
		:	
Db	361	LESVKFASHCLMNHPDLAKDLTEYLKVRSHDPEEAI RHDVIVTIITAGKRDLVNDQLL	420
Qy	361	GFVRERTLDKRWRVRKEAMMGLAQLYKKYCLHGEAGKEAAEKVSWIKDKLLHIYYQNSID	420
		:	
Db	421	GFVRERTLDKRWRVRKEAMMGLAQLYKKYCLHAEAGKDAAEKVSWIKDKLLHIYYQNSID	480
Qy	421	DKLLVEKIFAQYLVPHNLETEERMKCLYYLYASLDPNVAVKALNEMWKCQNMLRSHVRELL	480

Db	481	DKLLVEKIFAQYLVPHNLETEERMKCLYLYASLDPNAVKALNEMWKCQNMLRSHVRELL	540
Qy	481	DLHKQPTSEANCSAMFGKLMITIAKNLPDPGKAQDFVKKFNQVLGDDEKLRSQLLELLISPT	540
Db	541	DLHKQPTSEANSAAMFGKLMITIAKNLPDPGKAQDFVKKFNQVLGDDEKLRSQLLELLISPT	600
Qy	541	CSCCKQADICVREIARKLANPKQPTNPFLEMVKFLLERIAPVHIDSEAI SALVKLMNKSIE	600
Db	601	CSCCKQADVCVREIARKLANPKQPTNPFLEMVKFLLERIAPVHIDSEAI SALVKLMNKSIE	660
Qy	601	GTADDEEEGVSPDTAIRSGLELLKVLSTHPTSFHSAETYESLLQCLRMEDEKVAEAAIQ	660
Db	661	GTADDEEEGVSPDTAIRAGLELLKVLSTHPTSFHSAETYESLLQCLRMEDEKVAEAAIQ	720
Qy	661	IFRNTGHKIEIDLPIRSTLIPILHQAQKRGTPHQAQAVHCHAIFTNKEVQLAQIFEP	720
Db	721	IFRNTGHKIEIDLPIRSTLIPILHQAQKRGTPHQAQAVHCHAI FSNKEVQLAQIFEP	780
Qy	721	LSRSLNADVPEQLITPLVSLGHISMLAPDQFASPMKSVVANFIVKDLLMNDRSTGEKNGK	780
Db	781	LSRSLNADVPEQLITPLVSLGHISMLAPDQFASPMKSVVANFVVKDLLMNDRSTGEKNGK	840
Qy	781	LWSPDEEVSPEVLAKVQAIKLLVRWLLGMKNQSKSANSTLRLLSAMLVSEGLDTEQKRI	840
Db	841	LWSPDEEVSPEVLAKVQAIKLLVRWLLGMKNQSKSANSTLRLLSAMLVSEGLDTEQKRI	900
Qy	841	SKSDMSRLRLAAGSAIMKLAQEPYHEIITPEQFQLCALVINDECYQVRQIFAQKLHKAL	900
Db	901	SKSDMSRLRLAAGSAIMKLAQEPYHEIITPEQFQLCALVINDECYQVRQIFAQKLHKAL	960
Qy	901	VKLLLPLEYMAIFALCAKDPVKERRAHARQCLLKNI SIRREYIKQNPMATEKLLSLLPEY	960
Db	961	VKLLLPLEYMAIFALCAKDPVKERRAHARQCLLKNI SIRREYIKQNPMANEKLLSLLPEY	1020
Qy	961	VVPYMIHLLAHDPDFTRSQDQDLRDIKECLWFMLEVLMTKNENNSHAFMKKMAENIKLT	1020
Db	1021	VVPYMIHLLAHDPDFTKPDQDQDLRDVKECLWFMLEVLMTKNENNSHAFMKKMAENIKLT	1080
Qy	1021	RDAQSPDESKTNEKLYTVCDAVLCVINSKSALCNADSPKDPVLPKFFFTQPEKDFCNDKS	1080
Db	1081	RDAQSPDEPKANEKLYTVCDAVLCVINSKSALCNADSPKDPVLPKFFFTQPEKDFSNDRN	1140
Qy	1081	YISEETRVLLLTGKPKPAGVLGAVNKPLSATGRKPYVRSTGTETGSNINVSELNPSTGN	1140
Db	1141	YISEETRVLLLTGKPKPTGVLDTVNKPLSATGRRPYIRTTGSETGSNISVSELSSSAGN	1200
Qy	1141	RSREQSSEAAETGVSENEENPVRIISVTPVKNI DPVNKEINSDQATQGNISSDRGKKRT	1200
Db	1201	RSREQSSDISETGVSENDENPVRIISVTPAKT-EPVKNKEINSDQATQGN-STERGKKRT	1258
Qy	1201	VTAAGAENIQKTDK-EVDKESGPPAPSKPRRGRRPKSESQGNATKNDDLNPINKGRKRA	1259
Db	1259	ATASGTENIHQKAEENNADETGPSLAAKTRRGRPPKPEPQGTAKNEETNKPVRGRKRA	1318
Qy	1260	AVGQESPGGLEAGNAKAPKLQDLAKKAAPAERQIDLQR	1297
Db	1319	AASQESPGSLEAGNAKAPKOOTAKKPAQAQRIIDLQR	1356

ID Q4KLU7\_XENLA PRELIMINARY; PRT; 1323 AA.  
AC Q4KLU7;  
DT 02-AUG-2005, integrated into UniProtKB/TrEMBL.  
DT 02-AUG-2005, sequence version 1.  
DT 07-FEB-2006, entry version 2.  
DE Hypothetical protein.  
OS *Xenopus laevis* (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
OC Xenopodinae; Xenopus; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Oocytes;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for *Xenopus* research: The NIH *Xenopus*  
RT initiative.";  
RL Dev. Dyn. 225:384-391(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Oocytes;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Oocytes;  
RA Klein S., Gerhard D.S.;  
RL Submitted (JUL-2005) to the EMBL/GenBank/DDBJ databases.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; BC098992; AAH98992.1; -; mRNA.  
KW Hypothetical protein.  
SQ SEQUENCE 1323 AA; 149527 MW; CC1C1F80CFD0FFBD CRC64;

Query Match 88.0%; Score 5851; DB 2; Length 1323;  
Best Local Similarity 87.6%; Pred. No. 5.8e-307;  
Matches 1138; Conservative 73; Mismatches 78; Indels 10; Gaps 7;

